APPLICATIONS OF STEIN'S METHOD AND LARGE DEVIATIONS PRINCIPLE'S IN MEAN-FIELD $O(N)$ MODELS

BY

TAYYAB NAWAZ

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Mathematics
with a concentration in Computational Science and Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2017

Urbana, Illinois

Doctoral Committee:

Associate Professor Lee DeVille, Chair
Associate Professor Kay Kirkpatrick, Director of Research
Associate Professor Zoi Rapti
Associate Professor Anil Hirani

Abstract

In the first part of this thesis, we will discuss the classical XY model on complete graph in the mean-field (infinite-vertex) limit. Using theory of large deviations and Stein's method, in particular, Cramér and Sanov-type results, we present a number of results coming from the limit theorems with rates of convergence, and phase transition behavior for classical XY model. In the second part, we will generalize our results to mean-field classical N-vector models, for integers $N \geq 2$. We will use the theory of large deviations and Stein's method to study the total spin and its typical behavior, specifically obtaining non-normal limit theorems at the critical temperatures and central limit theorems away from criticality. Some of the important special cases of these models are the XY ($N = 2$) model of superconductors, the Heisenberg ($N = 3$) model (previously studied in [KM13] but with a correction to the critical distribution here), and the Toy ($N = 4$) model of the Higgs sector in particle physics.

Acknowledgments

First of all, I would like to express my sincere and deepest gratitude to my advisor, Professor Kay Kirkpatrick. I would like to thank you for supervision, support, and encouragement, which gave me self-assurance to broaden my research interests and helped me to grow as a mathematician.

I would like to thank Professor Richard S. Laugesen, the director of graduate studies, for his continuous advise, support and encouragement to explore various directions of mathematics throughout my PhD program at Illinois. I would also like to thank Professor Lee DeVille, Professor Zoi Rapti and Professor Anil Hirani for serving on my thesis committee. I would also like to thank the faculty and staff of the department of mathematics at Illinois for their constant help and support for last several years. I am also very appreciative of all the financial support I received over the years: National Science Foundation (NSF) CAREER award DMS-1254791, and NSF Grant 0932078 000 while in residence at Mathematical Sciences Research Institute (MSRI) during Fall 2015, Research Experience for Graduate Students through NetMath program for funding my Spring 2017 semester which enabled me to focus more on my research and complete my dissertation in a timely manner.

Finally, I would like to send my sincere gratitude to my whole family especially to my mother and my wife who constantly stood behind me and knew that I would succeed. Their dedication, support and encouragement was in the end what made this dissertation possible.

Table of Contents

List of Symbols

$\mathbb{S}^{N-1}$	N-dimensional unit hypersphere
σ	A spin configuration consisting of n spins in $\mathbb{S}^{N-1}$
Ω_n	the set of all possible spin configurations for n spins in $\mathbb{S}^{N-1}$
S_n	the total spin for a spin configuration σ
M_n	the average spin for a spin configuration σ
β	non-negative inverse temperature
I_β	the rate function for the non-negative inverse temperature β
$\mathbb{R}$	the set of real numbers
Z	the partition function
φ	the free energy functional
d_{BL}	bounded Lipschitz distance between two random variables
T_p	an invertible characterizing operator
$H(\cdot \mid \mu)$	the relative entropy for a probability measure ν with respect to uniform measure μ

List of Figures

Chapter 1

Introduction

In statistical mechanics, mean-field models are often the starting point for understanding the behavior of the underlying physical systems, at least in high dimensions. In particular, we can use large deviations to study the asymptotics of physical quantities such as magnetization (total spin in our terminology) in models such as XY for superconductors, Heisenberg for ferromagnets [KM13], or the Toy model for particle physics. There is a family of models generalizing these important special cases, namely the mean-field classical N-vector spin models, where each spin σ_i is in an N-dimensional unit hypersphere, at a lattice site or in our case at a complete graph vertex i among n vertices. Then in the absence of an external field, each microstate or spin configuation $\sigma = (\sigma_1, \sigma_2, ..., \sigma_n)$ in the configuration space $\Omega_n = (\mathbb{S}^{N-1})^n$ has a Hamiltonian defined by:

$$H_n(\sigma) = -\sum_{i,j} J_{i,j} \langle\ \sigma_i, \sigma_j \rangle.$$

where $|\sigma_i| = 1$ and θ_i is the angle corresponding to i-th spin which it makes with some arbitrary axis. For the mean-field models defined on the complete graph, every two vertices (i, j) are adjacent and the interaction between σ_i and σ_j is given by the constant $J_{i,j} = \frac{1}{2n}$, which can be viewed as an averaged interaction.

The simplest spin model is the Ising model, with one-dimensional ± 1 spins, a model that is used extensively in statistical mechanics, invented by Ernest Ising while working with his advisor Wilhelm Lenz [Isi25, Bru67]. The one-dimensional Ising model has no phase transition, but there is a phase transition on an infinite two-dimensional lattice. Near the critical temperature, Landau argued, we can express the free energy as a Taylor expansion in the order parameter. At zero-field magnetization, we can observe a qualitative change in behavior at some critical temperature to see the phase transition [Gri64].

The $N = 1$ case of the mean-field N-vector model is the Curie-Weiss model, which approximates the Ising model well for higher dimensions. The normalized total spin in the Curie-Weiss model has a non-Gaussian law in the non-critical regimes, and a law that converges to the distribution with density proportional to $e^{-x^4/12}$ at the critical temperature (Ellis and Newman [EN78a]). In the absence of an external field certain complex systems are naturally attracted to critical points, exhibiting the phenomenon of self-organized criticality [BTW87]. Chatterjee and Shao proved that the total spin at the critical temperature for this non-central limit theorem satisfies an error bound of order $\frac{1}{\sqrt{n}}$ [CS11]. For the N-vector Curie-Weiss model in the absence of external field the average magnetization is used for a complete description of the equilibrium states [AZ85]. Verbeure and Zagrebnov used the Laplace method to derive the probability distribution in a decaying external field as well as in the absence of an external field [VZ94]. Later, it was proved that the free energy and the equilibrium state of the N-vector model on a finite lattice allows complete asymptotic expansions in powers of $\frac{1}{n}$ [AVZ00].

The XY model, with two-dimensional circular spins, models superconductors and is interesting but challenging to study its phase transition rigorously [And58]. For instance, the Mermin-Wagner theorem states that in two spatial dimensions, such a continuous symmetry cannot be broken spontaneously at any finite temperature [MW66]. Thus the 2D XY model cannot have an ordered phase at low temperature like the Ising model does. Stanley and Moore provided some evidence that this system has a phase transition but it can't be of usual type with finite mean magnetization below the critical temperature [Sta68, Moo69]. But still in two dimensions it has an infinite-order phase transition named after Berezinski, Kosterlitz and Thouless (BKT phase transition), who discovered in Nobel Prize winning work that there is phase transition from bound vortex-antivortex pairs at low temperature to unpaired vortices and anti-vortices at some critical temperature [KT73]. Above the transition temperature correlations between spins decay exponentially as usual, with some correlation length. They also showed that this system does not have any long-range order as the ground state is unstable against low-energy spin-wave excitations. They further proved that this is a low-temperature quasi-ordered phase with a correlation function that decreases with the distance like a power, which depends on the temperature.

Because the two-dimensional lattice XY model is challenging, the mean-field case is often the starting point for rigorous analysis of these spin models, a case that can be thought of as a large-dimensional limit of nearest-neighbor lattice models, or as an infinite limit for complete graph models. It is known that the classical N-vector model with spins in $\mathbb{S}^{N-1}$, in the large-dimensional ($d \to \infty$) limit on the lattice $\mathbb{Z}^d$, has the critical inverse temperature $\beta_c = N$ [KS89]. This limiting case is thought to approximate high-dimensional models well because magnetization goes to zero below the critical temperature for all d, and the magnetization goes to the correct limit above the critical temperature as $d \to \infty$.

In this thesis, we study mean-field N-vector models for positive integers N, in the infinite limit for complete graphs. Using the theory of large deviations along with Stein's method-type limit theorems, we describe the asymptotic behavior of the $O(N)$ spin models [CS11, KM13, KN16a]. The material discussed in chapter 2 is from the paper [KN16b], while chapters 3-5 are based on the papers [KN16a, KN16b]. This thesis is organized as follows: In chapter 2, we discuss the asymptotic results related to classical XY model on a complete graph in the mean-field limit. We discuss Large deviation principles (LDPs) for the total spin (with rate functions) and the empirical spin distribution (with relative entropies) for each non-negative inverse temperature. We also derived the limit theorems for total spin in each phase. In chapter 3, we generalized these asymptotic results to mean-field classical N-vector models, for integers $N \geq 2$. In particular, we derived a non-normal limit theorem for the total spin at the critical temperature, with limiting density of the squared length proportional to $t^{\frac{N-2}{2}} e^{-\frac{1}{4N^2(N+2)} t^2}$. Chapters 4-5 contains some technical details including calculus for the free energy, and abstract results for the non-normal Stein's method application. Lastly, chapter 6 contains details of our current project which is related to the computational aspects of mean field $O(N)$ models.

Chapter 2

Asymptotic behavior of the mean-field XY model

We use mean-field theory to approximate a challenging problem and to study a many-body problem by converting it into a one-body problem. In this chapter, mostly taken from [KN16b], we describe the asymptotic results for the XY model mostly without proofs. While in next chapter, we will give the proofs for the N-vector model, therefore the proofs for the XY case are straightforward exercises with $N = 2$.

The XY model on a complete graph K_n with n vertices in the absence of an external field is defined as follows: there is a circular spin $\sigma_i \in \mathbb{S}^1$ at each site $i \in 1, 2, ..., n$. The configuration space of the XY model is $\Omega_n = (\mathbb{S}^1)^n$ where each microstate is $\sigma = (\sigma_1, \sigma_2, ..., \sigma_n)$. The corresponding Hamiltonian energy is defined by

$$H_n(\sigma) = -\sum_{i,j} J_{i,j} \langle\ \sigma_i, \sigma_j \rangle$$

2.1 Mean-field XY model

We consider the isotropic mean-field classical XY model on a finite complete graph K_n with n vertices. That is, at each site $i \in K_n$ of the graph is a spin living in $\Omega = \mathbb{S}^1$, so the state space is $\Omega_n = (\mathbb{S}^1)^n$. See Fig. 2.1 for a picture of the XY model on 5 vertices. The mean-field interaction for the XY model on the complete graph is defined by $J_{i,j} = \frac{1}{2n}$ for all i, j.

The corresponding mean-field Hamiltonian energy $H_n : \Omega_n \to \mathbb{R}$ is given by:

$$H_n(\sigma) := -\frac{1}{2n} \sum_{i,j=1}^{n} \langle\ \sigma_i, \sigma_j \rangle = -\frac{1}{2n} \sum_{i,j} \cos(\theta_i - \theta_j),$$

where θ_i is the angle that the i-th spin makes with respect to some axis. The corresponding Gibbs measure is the probability measure $P_{n,\beta}$ on Ω_n with density function:

$$f(\sigma) := \frac{1}{Z(\beta)} e^{-\beta H_n(\sigma)}. \tag{2.1}$$

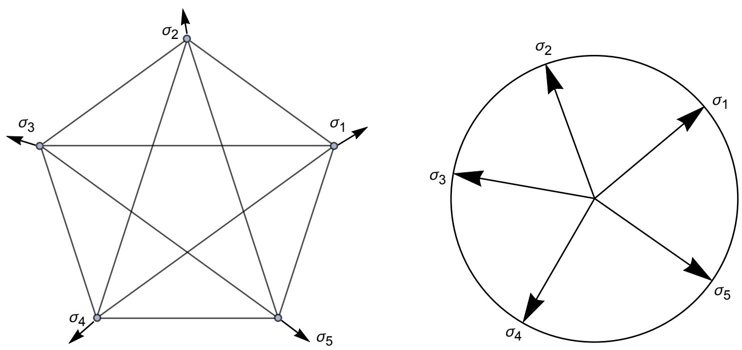

Figure 2.1: Left: The classical mean-field XY Model on the complete graph K_5 with five sample spin vectors. Right: The projection of the same spin vectors from K_5 onto $\mathbb{S}^1$.

where Z is the normalizing constant, also known as the partition function, which encodes the statistical properties of the model such as free energy and magnetization. Note that Gibbs measure here is a normalization of the Boltzmann distribution, and that the inverse temperature β is equal to $(k_B T)^{-1}$, where k_B is the Boltzmann constant and T is the temperature of the system. We can understand the structural behavior of the spin vectors distribution by studying extreme cases for the inverse temperature β as follows:

- At high temperature, from equation (2.1) we can predict that the Gibbs measure is uniform.
- At low temperature, again from equation (2.1) we can predict that the Gibbs measure decays quickly, and the spin vector distribution prefers the lowest-energy ground state.

The most likely physical system states corresponding to the Gibbs measure are called the canonical macrostates. We will consider the random measure of the spins $\{\sigma_i\}$, defined by

$$\mu_{n,\sigma} := \frac{1}{n}\sum_{i=1}^{n} \delta_{\sigma_i}$$

on $\mathbb{S}^1$ and study the total empirical spin, defined by $S_n(\sigma) := \sum_{i=1}^{n} \sigma_i$.

The relative entropy of a probability measure ν on $\mathbb{S}^1$, with respect to the uniform probability measure μ is defined by

$$H(\nu \mid \mu) := \begin{cases} \int_{\mathbb{S}^1} f \log(f) d\mu & if\ f := \frac{d\nu}{d\mu}\ exists; \\ \infty & otherwise. \end{cases} \tag{2.2}$$

2.2 LDPs, Free Energy and Macrostates

Let $M_1(\mathbb{S}^1)$ represent the probability measures on $\mathbb{S}^1$ with the weak-$*$ topology. We are interested in analyzing the total spin, $S_n := \sum_{i=1}^n \sigma_i$, as a function of the inverse temperature β in the Gibbs measures. This leads us to consider large deviation principles (LDPs) for the $\mu_{n,\sigma}$, and then we can rewrite the free energy more explicitly to describe the phase transition at $\beta = 2$. Part of Theorem 2.2.1 ($\beta = 0$) is a special case of Sanov's theorem, while the other cases ($\beta > 0$) follow from an abstract result of Ellis, Haven, and Turkington ([EHT00], Theorem 2.5).

Theorem 2.2.1. *If P_n is the n-fold product of uniform measure on $\mathbb{S}^1$ and $\mu_{n,\sigma}$ is the random measure as defined above. For $\Gamma \subset M_1(\mathbb{S}^1)$, the $\mu_{n,\sigma}$ satisfy an LDP with rate function*

$$I_\beta(\nu) := H(\nu \mid \mu) - \frac{\beta}{2}\left|\int_{\mathbb{S}^1} x d\nu(x)\right|^2 - \varphi(\beta), \tag{2.3}$$

where the free energy is given by

$$\varphi(\beta) = \inf_{\nu \in M_1(\mathbb{S}^1)} \left[H(\nu \mid \mu) - \frac{\beta}{2}\left|\int_{\mathbb{S}^1} x d\nu(x)\right|^2\right]. \tag{2.4}$$

For fixed $\beta \geq 0$, every subsequence of $P_{n,\beta}\left[\mu_{n,\sigma} \in \cdot\right]$ converges weakly to a probability measure on $M_1(\mathbb{S}^1)$ concentrated on the canonical macrostates $\mathcal{E}_\beta := \{\nu : I_\beta(\nu) = 0\}$, i.e., the zeros of the rate function.

For $\beta = 0$, the relative entropy $H(\cdot \mid \mu)$ achieves its minima of 0 only for the uniform measure μ, implying that the canonical macrostate is disordered. For $\beta > 0$, canonical macrostates are defined abstractly through zeros of the rate function (2.3), and later Theorem 2.2.5 will describe the macrostates explicitly.

The free energy given by (2.4) can be transformed into the following more explicit form.

Theorem 2.2.2. *The free energy φ has the formula:*

$$\varphi(\beta) = \begin{cases} 0, & \text{if } \beta < 2, \\ \Phi_\beta(g^{-1}(\beta)), & \text{if } \beta \geq 2, \end{cases}$$

where I_i below are modified Bessel functions of first kind and Φ_β is the functional defined by:

$$\Phi_\beta(r) := r\frac{I_1(r)}{I_0(r)} - \log\left[I_0(r)\right] - \frac{\beta}{2}\left(\frac{I_1(r)}{I_0(r)}\right)^2, \tag{2.5}$$

and

$$g(r) := r\frac{I_0(r)}{I_1(r)}.$$

Here the phase transition is continuous as the function φ and its derivative φ' are continuous at the critical threshold $\beta = \beta_c = 2$.

We can prove the statement about order of phase transition using calculus of one dimensional function φ' as follows:

$$\lim_{\beta\to 2^+} \varphi'(\beta) = \lim_{\beta\to 2^+} \Phi'_\beta(g^{-1}(\beta))(g^{-1})'(\beta).$$

Now using the inverse function theorem for $x > 0$, where $g'(x) > 0$, we have

$$\lim_{\beta\to 2^+} \varphi'(\beta) = \lim_{\beta\to 2^+} \frac{(xI_0(x) - \beta I_1(x))\left(I_0(x)^2 - 2I_1(x)^2 + I_0(x)I_2(x)\right)}{2I_0(x)^3}\left(\frac{1}{g'(x)}\right),$$

Using the definition of $g(r)$ we conclude that

$$\lim_{\beta\to 2^+} \varphi'(\beta) = 0.$$

The magnetization for the classical XY model can be obtained by differentiating the partition function:

$$|m| = \left|\mathbb{E}\left[\frac{1}{n}\sum_i \sigma_i\right]\right| = \left|\mathbb{E}\left[\frac{1}{n}S_n\right]\right| = \frac{I_1(r)}{I_0(r)}$$

From the free energy we can precisely explain the phase transition at $\beta = 2$. For $0 \le \beta \le 2$, we have a unique global minimum for the free energy at the origin with a zero magnetization. For $\beta \ge 2$, we have a unique global minimum for a positive radius.

Let $\{\sigma_i\}_{i=1}^n$ be i.i.d. uniform random points on $\mathbb{S}^1 \subseteq \mathbb{R}^2$. We have the following Cramér-type LDP for the average spin. $M_n := \frac{1}{n}\sum_{i=1}^n \sigma_i$.

Theorem 2.2.3. *Let $P_{n,\beta}$ be the Gibbs measure defined above* (2.1). *Then for $\beta \ge 0$, the average spin $M_n = M_n(\sigma) := \frac{1}{n}\sum_{i=1}^n \sigma_i$ satisfies an LDP with rate function $I_\beta(x) = \Phi_\beta(r)$:*

$$P_{n,\beta}\left(M_n \simeq x\right) \simeq e^{-n\Phi_\beta(r)},$$

where Φ_β is given by (2.5) and $r = |x|$.

For an explicit representation of $\mathcal{E}_\beta$, we note from (2.2) that the relative entropy depends only on the distribution of f. By Fubini's theorem

$$\int f\log(f)d\mu = \int_0^\infty \mu\big[f\log(f) > t\big]dt - \int_0^\infty \mu\big[f\log(f) < -t\big]dt.$$

This implies that for a fixed f, the quantity $\left|\int x d\nu(x)\right|$ is maximized for corresponding densities which are symmetric about a fixed pole and decreasing as the distance from the pole increases. Using this reasoning, consider a density f that is symmetric about the north pole and decreasing away from the pole i.e., ν_f is the measure with density $f(x,y) = f(y)$ which is increasing in y. Then

$$\begin{aligned} H(\nu_f \mid \mu) &= \frac{1}{2\pi}\int_{\mathbb{S}^1} f(x,y)\log[f(x,y)]dxdy \\ &= \frac{1}{2\pi}\int_0^{2\pi} f(\cos(\theta))\log[f(\cos(\theta))]d\theta \\ &= \frac{1}{\pi}\int_{-1}^{1} \frac{f(y)\log[f(y)]}{\sqrt{1-y^2}}dy. \end{aligned}$$

Similarly,

$$\int_{\mathbb{S}^1} x d\nu_f(x) = \frac{1}{\pi}\begin{bmatrix} 0 \\ 1 \end{bmatrix}\int_{-1}^{1}\frac{yf(y)}{\sqrt{1-y^2}}dy.$$

Therefore, our minimization problem is reduced to minimizing the following functional

$$\frac{1}{\pi}\int_{-1}^{1}\frac{f(y)\log[f(y)]}{\sqrt{1-y^2}}dy - \frac{\beta}{2}\left(\frac{1}{\pi}\int_{-1}^{1}\frac{yf(y)}{\sqrt{1-y^2}}dy\right)^2$$

over $f : [-1,1] \to \mathbb{R}_+$ such that f is increasing and $\frac{1}{\pi}\int_{-1}^{1}\frac{f(y)}{\sqrt{1-y^2}}dy = 1$. We can rewrite the first term of the last expression to see that it involves the usual entropy $S(f) = \int f\log(f)$:

$$\frac{1}{\pi}\int_{-1}^{1}\frac{f(y)\log[f(y)]}{\sqrt{1-y^2}}dy = \frac{1}{\pi}\int_{-1}^{1}\frac{f(y)}{\sqrt{1-y^2}}\log\left[\frac{f(y)}{\pi}\right]dy + \log(\pi) = -S\left(\frac{f}{\pi}\right) + \log(\pi).$$

Now for $\left|\int x d\nu(x)\right| = c \in [0,1]$, using constrained entropy maximization (see Theorem 12.1.1 from [CT06]), we will minimize $\frac{1}{\pi}\int_{-1}^{1}\frac{yf(y)}{\sqrt{1-y^2}}dy$, that is, maximize $S(f/\pi)$, over the $\nu \in M_1(\mathbb{S}^1)$ corresponding to this c.

Proposition 2.2.4. *Consider a set of functions $f : [-1,1] \to \mathbb{R}_+$, with weight function $w(y) = \frac{1}{\sqrt{1-y^2}}$, such that $\int_{-1}^{1} f(y)w(y)dy = 1$, and $\left|\int_{-1}^{1} yf(y)w(y)dy\right| = c$. i.e., weighted integral of f is 1 while first weighted moment is bounded. Then the exponential function $f^*(y) = \pi a e^{by}$ uniquely maximizes $S(f/\pi)$ over the densities satisfying these conditions.*

For $c \in [0,1]$, observe that f^* increasing gives all $b \in [0,\infty)$. Now for $b \in [0,\infty)$, our functional minimization reduces to the following one dimensional function:

$$\frac{1}{\pi}\int_{-1}^{1} \frac{f(y)\log[f(y)]}{\sqrt{1-y^2}}dy - \frac{\beta}{2}c^2 = b\frac{I_1(b)}{I_0(b)} - \log\left[I_0(b)\right] - \frac{\beta}{2}\left(\frac{I_1(b)}{I_0(b)}\right)^2$$
$$=: \Phi_\beta(b).$$

The following theorem, a special case proved using the calculus of variations in [KN16a], describes the canonical macrostates:

Theorem 2.2.5. *1. For $\beta \leq 2$, $\inf_{b\geq 0}\{\Phi_\beta(b)\} = 0$ is achieved for $b = 0$ and the corresponding $a = 1$, so that the minimizing function $f^* = 1$ and therefore the only canonical macrostate is the uniform distribution μ.*

2. For $\beta > 2$, $\inf_{b\geq 0}\{\Phi_\beta(b)\} = \Phi_\beta(g^{-1}(\beta))$, where $b = g^{-1}(\beta)$ is the unique strictly positive solution to $g(b) = \beta$ where

$$g(b) = b\frac{I_0(b)}{I_1(b)},$$

$a = \frac{1}{\pi I_0(b)}$ and $\lim_{\beta\downarrow 2}\inf_{b\geq 0}\{\Phi_\beta(b)\} = 0$. In this case, the canonical macrostates are given by $\mathcal{E}_\beta = \{\nu_{f,x}\}_{x\in\mathbb{S}^1}$, where $\nu_{f,x}$ is the measure that is the rotation of ν_f from north pole to x-direction, which is symmetric about the north pole with density $f : [-1,1] \to \mathbb{R}$ given by $f(y) = \pi a e^{by}$ with a and b as above.

We can also visualize the Gibbs measure corresponding to subcritical or supercritical cases as shown in Fig. 2.2.

2.3 Limit Theorems for the Total Spin

Next we understand the asymptotics for the total spin of the mean-field XY model, in different regimes across the phase transition, describing the central and non-central limit theorems for each phase.

In the high temperature regime ($0 \leq \beta < 2$), the average spin (magnetization) of the system goes to zero with increasing number of spins $n \to \infty$, and we have a multivariate central limit theorem with a rate of convergence in Theorem 2.3.1. The main idea is to use Stein's method [KM13, Ste86, Mec09] with the exchangeable pair (W_n, W_n') from the Gibbs sampling approach: our random variable representing the

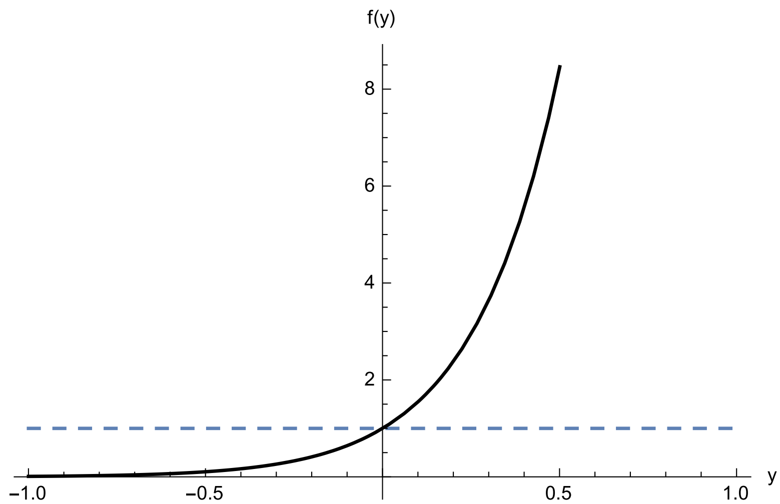

Figure 2.2: Cross-sections of two canonical macrostates: For $\beta \leq 2$ (the disordered regime), we have the uniform distribution $f(y) = 1$ as the dotted line; for $\beta = 5 > \beta_c = 2$ (the ordered regime), we have plotted the cross-section of the distribution ν_f, given by $f(x, y) = f(y) = \frac{e^{by}}{I_0(y)}$, showing that the spins point predominantly near the north-pole direction.

rescaled total spin of the original configuration is

$$W_n := \sqrt{\frac{2-\beta}{n}} \sum_{i=1}^{n} \sigma_i,$$

while the random variable representing the rescaled total spin of the new configuration, with $I \in \{1, \dots, n\}$ chosen uniformly at random, is

$$W_n' := W_n(\sigma') = W_n - \sqrt{\frac{2-\beta}{n}} \sigma_I + \sqrt{\frac{2-\beta}{n}} \sigma_I'.$$

Our multivariate central limit theorem in the high temperature regime is as follows:

Theorem 2.3.1. *In the high temperature regime $0 < \beta < 2$, if W_n is defined as above, Z is a standard normal random variable in $\mathbb{R}^2$, c_β is a function depending on β only, $L(g)$ is the modulus of uniform continuity of g, and $M(g)$ is the maximum operator norm of the Hessian of g, then we have:*

$$\sup_{g: L(g), M(g) \leq 1} |\mathbb{E}\, g(W_n) - \mathbb{E}\, g(Z)| \leq \frac{c_\beta}{\sqrt{n}}$$

The proof of Theorem 2.3.1 proceeds in several steps, as a special case of [KN16a]: first we use the fact that the density of the Gibbs measure is rotationally invariant to conclude that each spin has a uniform marginal distribution. We obtain the complete asymptotic behavior of the total spin using the rotational invariance of the total spin, a strategy adapted from [KM13]. We calculate the variance of the total spin to arrive at the proper scaling for defining the exchangable pair and use the pair to derive expressions and bounds for the linear factor Λ appearing in the conditional expectation and the remainder terms R and R' [KM13, KN16a, Mec09]. The rest follows from a theorem of Meckes [Mec09].

As the temperature decreases to zero, the spins start aligning. For smaller values of $\beta > 2$, the spins vectors are aligned weakly, while for larger β, this alignment is strong. For any $\beta > 2$, because of the large deviation principle in Theorem 2.2.3, we have that $|\sum \sigma_j|$ is close to bn/β with high probability, if b is the minimizer in Φ_β. And due to the circular symmetry, all points on the circle of radius bn/β are equally likely. With this reasoning, similar to [KM13], it is natural to consider the random variable representing the fluctuations of squared-length of total spin, i.e.,

$$W_n := \sqrt{n}\left[\frac{\beta^2}{n^2 b^2}\left|\sum_{j=1}^{n}\sigma_j\right|^2 - 1\right]. \tag{2.6}$$

Our multivariate central limit theorem in the low temperature (ordered) regime is as follows:

Theorem 2.3.2. *If $\beta > 2$ and b is the solution of $b = \beta f(b) := \beta\frac{I_1(b)}{I_0(b)}$, and W_n is as defined above in* (2.6)*, and if Z is a centered normal random variable with variance V, where*

$$V = \frac{4\beta^2}{(1-\beta f'(b))\, b^2}\left[1 - \frac{1}{b}\frac{I_1(b)}{I_0(b)} - \left(\frac{I_1(b)}{I_0(b)}\right)^2\right],$$

then there exists c_β, depending only on β, then

$$d_{BL}(W_n, Z) \le c_\beta\left(\frac{\log(n)}{n}\right)^{1/4}.$$

where $d_{BL}(X,Y)$ is the bounded Lipschitz distance between random variables X and Y.

Again the proof of Theorem 2.3.2 follows from a univariate analogue of the abstract normal approximation of Stein [Ste86], and relies on conditional moment bounds. The fact that the variance is positive was proved by Amos [Amo74b] while deriving the improved bounds on the ratio of Bessel functions.

At the critical temperature $\beta_c = 2$, we will consider the random variable

$$W_n := \frac{c}{n^{3/2}}\sum_{i,j=1}^{n}\langle\, \sigma_i, \sigma_j\rangle, \tag{2.7}$$

and make an exchangeable pair (W_n, W_n') using Glauber dynamics. Using symmetry of the total spin and

Stein's method similar to [CS11, KM13], we will obtain critical limiting density function p as defined below.

Theorem 2.3.3. *For the critical inverse temperature $\beta = 2$, if W_n is as defined above in (2.7), and X is the random variable with the density*

$$p(t) = \begin{cases} \frac{1}{Z} e^{-t^2/64} & t \geq 0, \\ 0 & t < 0, \end{cases}$$

where Z is normalizing constant, then there exists a universal constant C such that

$$\sup_{\substack{\|h\|_\infty \leq 1, \|h'\|_\infty \leq 1 \\ \|h''\|_\infty \leq 1}} \left| \mathbb{E}\, h(W_n) - \mathbb{E}\, h(X) \right| \leq \frac{C \log(n)}{\sqrt{n}}.$$

The proof of the limit theorem for the critical temperature is essentially via the "density approach" to Stein's method introduced by Stein, Diaconis, Holmes, and Reinert [SDHR04]. Recenlty, also Chatterjee and Shao [CS11] have applied this approach to the total spin of the mean-field Ising model, i.e., the Curie-Weiss model.

We note that these limit theorems with explicit rates of convergence can be generalized to high-dimensional spins, but we will omit those technicalities in the next chapter.

Chapter 3

Asymptotics of mean-field $O(N)$-models

3.1 The Setup and Large Deviations

We consider, for N which is a fixed positive integer, the mean-field classical N-vector, or $O(N)$, model on a complete graph K_n with n vertices (these are isotropic models, meaning no external magnetic field). Most of the material in this chapter is taken from [KN16a]. At each site i on the graph is a spin σ_i in $\Omega = \mathbb{S}^{N-1}$, so the state (or configuration) space is $\Omega_n = (\mathbb{S}^{N-1})^n$ with product measure P_n from the uniform probability measure on $\mathbb{S}^{N-1}$. For these models the mean-field Hamiltonian energy $H_n : \Omega_n \to \mathbb{R}$ is defined by:

$$H_n(\sigma) := -\frac{1}{2n} \sum_{i,j=1}^{n} \langle\, \sigma_i, \sigma_j \rangle. \tag{3.1}$$

The energy per particle is $h_n(\sigma) = \frac{1}{n} H_n(\sigma)$, and the canonical ensemble, or Gibbs measure, is the probability measure $P_{n,\beta}$ on Ω_n with density (with respect to P_n):

$$f(\sigma) := \frac{1}{Z} e^{-\beta H_n(\sigma)} = \frac{1}{Z} \exp\left(\frac{\beta}{2n} \sum_{i,j=1}^{n} \langle\, \sigma_i, \sigma_j \rangle \right).$$

The (normalizing) partition function is given by:

$$Z = Z_n(\beta) = \int_{\Omega_n} \exp\left(\frac{\beta}{2n} \sum_{i,j=1}^{n} \langle\, \sigma_i, \sigma_j \rangle \right) dP_n.$$

We will call $M_1(\mathbb{S}^{N-1})$ the space of probability measures on $\mathbb{S}^{N-1}$ with the weak-* topology.

Now we are interested in studying the behavior of the important physical quantity of total spin $S_n := \sum_{i=1}^{n} \sigma_i$ in terms of the inverse temperature β, distributed according to the Gibbs measures. We will start this section by presenting a proposition stating the large deviation principle for the empirical spin distribution for non-interacting particles (disordered infinite-temperature case) $\beta = 0$, a proposition that is a special case of Sanov's theorem.

Proposition 3.1.1. *If P_n is the n-fold product of uniform measure on $\mathbb{S}^{N-1}$, $\mu_{n,\sigma} = \frac{1}{n}\sum_{i=1}^{n}\delta_{\sigma_i}$ is the empirical spin distribution, and Γ a Borel subset of $M_1(\mathbb{S}^{N-1})$, then*

$$
\begin{aligned}
-\inf_{\nu\in\Gamma^\circ} H(\nu \mid \mu) &\leq \liminf_{n\to\infty}\frac{1}{n}\log P_n[\mu_{n,\sigma}\in\Gamma] \\
&\leq \limsup_{n\to\infty}\frac{1}{n}\log P_n[\mu_{n,\sigma}\in\Gamma] \leq -\inf_{\nu\in\overline{\Gamma}} H(\nu \mid \mu);
\end{aligned}
$$

i.e., the random measures $\mu_{n,\sigma}$ satisfy an LDP with rate function $H(\cdot \mid \mu)$, the relative entropy, defined for a probability measure ν on $\mathbb{S}^{N-1}$, with respect to uniform measure μ, by:

$$
H(\nu \mid \mu) := \begin{cases} \int_{\mathbb{S}^{N-1}} f\log(f)d\mu & \text{if } f := \frac{d\nu}{d\mu} \text{ exists;} \\ \infty & \text{otherwise.} \end{cases}
$$

In particular, at infinite temperature $\beta = 0$, the unique minimizer of the rate function is the uniform measure μ, meaning that spins are uniformly and independently distributed on the sphere, with no preferred direction in this disordered phase.

Now we state the general case $\beta \geq 0$, which follows from abstract results of Ellis, Haven, and Turkington ([EHT00], Theorems 2.4 and 2.5):

Theorem 3.1.2. *If $\beta \geq 0$ then the empirical spin distributions $\mu_{n,\sigma}$ satisfy an LDP on $M_1(\mathbb{S}^{N-1})$ with rate function:*

$$
I_\beta(\nu) := H(\nu \mid \mu) - \frac{\beta}{2}\left|\int_{\mathbb{S}^{N-1}} x d\nu(x)\right|^2 - \varphi(\beta),
$$

where φ is the free energy defined by $\varphi(\beta) := -\lim_{n\to\infty}\frac{1}{n}\log Z_n(\beta)$, which exists and is given by the alternative formula:

$$
\varphi(\beta) = \inf_{\nu\in M_1(\mathbb{S}^{N-1})}\left[H(\nu \mid \mu) - \frac{\beta}{2}\left|\int_{\mathbb{S}^{N-1}} x d\nu(x)\right|^2\right]. \tag{3.2}
$$

We can calculate the minima in the expression of this rate function and verify that in the subcritical regime ($\beta < N$) there is a unique minimum, while in the supercritical regime there is a family of minima parametrized by $\mathbb{S}^{N-1}$. The free energy given by (3.2) can be written in the following more explicit form using a method like the one in the previous section. In particular, we have a Cramér-type LDP for the average spin $M_n := \frac{1}{n}\sum_{i=1}^{n}\sigma_i \in \mathbb{R}^N$, with rate function $I_{\beta,N}(x) = \Phi_{\beta,N}(r)$, defined below for $\beta \geq 0$ and $r = |x|$.

Remarks on notation:

1. Throughout this chapter we will write the rate function as I_β where the subscript is the real, non-

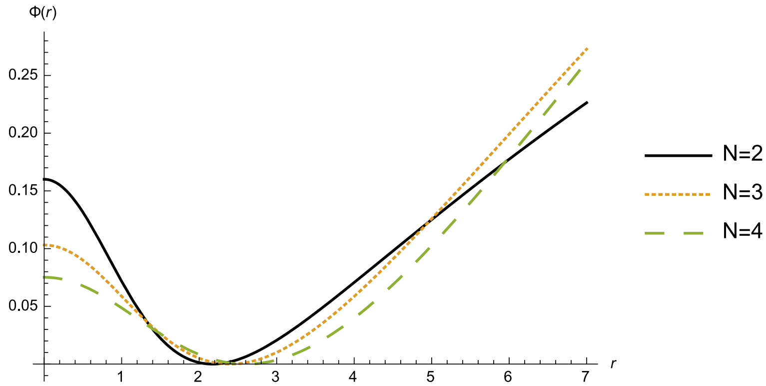

Figure 3.1: Graph of the rate function $I_{\beta,N}(x) = \Phi_{\beta,N}(r)$ in the supercritical regime ($\beta = N+1$) for $2 \leq N \leq 4$, which has minimum at radius $g_N^{-1}(\beta) = r$.

negative inverse temperature $\beta = 1/(k_B T)$, where T is temperature and k_B is the Boltzmann constant; we will write the modified Bessel function of the first kind as I_n where n is an integer.

2. Consider the $O(N)$ model with the above Hamiltonian (3.1), with N representing the dimension of the spin $\sigma_i \in \mathbb{S}^{N-1}$. Then on the complete graph K_n the $O(N)$ magnetization $M_{N,n} = \sum_{i=1}^{n} \sigma_i$ has the following mean-field limit:

$$|M_N| = \frac{I_{\frac{N}{2}}(r)}{I_{\frac{N}{2}-1}(r)}$$

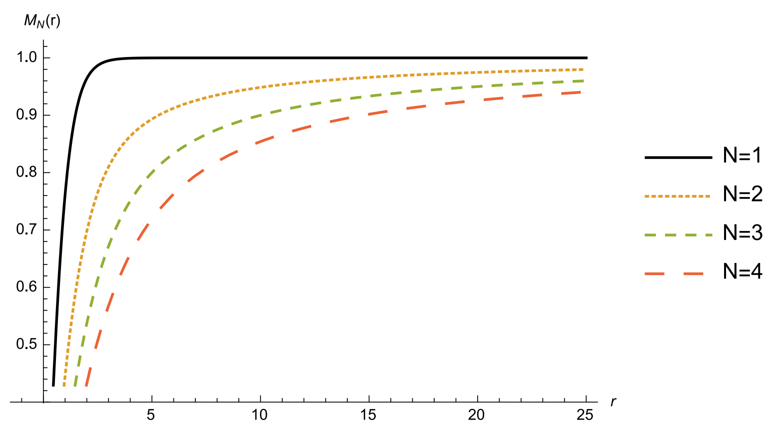

Figure 3.2: Graph of magnetization limits $|M_N|$ for N-vector models, $1 \leq N \leq 4$. For the mean-field Ising model, $M_1 = tanh(x)$, for the mean-field XY model $|M_2| = \frac{I_1(r)}{I_0(r)}$, for the mean-field Heisenberg model $|M_3| = coth(r) - \frac{1}{r}$, and for the mean-field Toy model of the Higgs sector, $|M_4| = \frac{I_2(r)}{I_1(r)}$. Here r and β are related by the formula $g_N(r) := r\frac{I_{\frac{N}{2}-1}(r)}{I_{\frac{N}{2}}(r)} = \beta$

3. Note that the rate functions I_β are always nonnegative, as illustrated in Figure 3.3 below for $N = 2$ in three representative cases of β.

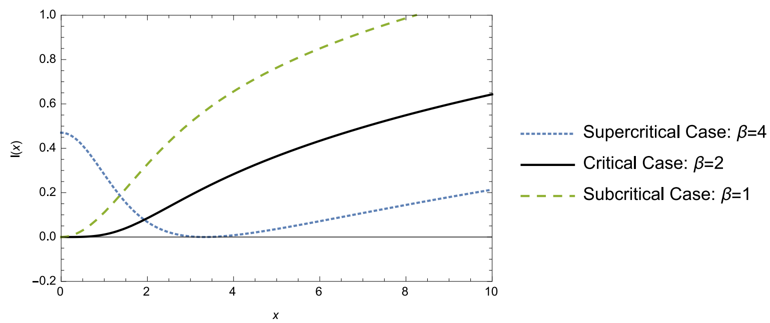

Figure 3.3: Rate Function I_β for the mean-field XY Model

Theorem 3.1.3. *For dimension N , the free energy φ has the formula:*

$$\varphi(\beta) = \begin{cases} 0, & \text{if } \beta < N, \\ \Phi_\beta(\Upsilon^{-1}(\beta)), & \text{if } \beta \geq N, \end{cases}$$

where

$$\Phi_\beta(r) = r\frac{I_{\frac{N}{2}}(r)}{I_{\frac{N}{2}-1}(r)} + \log\left[\frac{A_N}{A_{N-1}}\frac{r^{\frac{N}{2}-1}}{B_N \pi I_{\frac{N}{2}-1}(r)}\right] - \frac{\beta}{2}\left(\frac{I_{\frac{N}{2}}(r)}{I_{\frac{N}{2}-1}(r)}\right)^2,$$

with

$$A_N := \frac{2\pi^{\frac{N}{2}}}{\Gamma\left(\frac{N}{2}\right)}, \qquad \Upsilon(r) = \Upsilon_N(r) := r\frac{I_{\frac{N}{2}-1}(r)}{I_{\frac{N}{2}}(r)} = \beta,$$

and

$$B_N = \begin{cases} \prod_{k=0}^{\frac{N}{2}-1} |2k-1|, & \text{if } N \text{ even}, \\ \frac{2^{\frac{N}{2}-1}\Gamma\left(\frac{N-1}{2}\right)}{\sqrt{\pi}}, & \text{if } N \text{ odd}. \end{cases} \tag{3.3}$$

In particular, we find the critical threshold $\beta = N$, and we can check by calculating limits that φ and φ' are continuous, implying that the phase transition is continuous.

We can precisely describe the phase transition in free energy as follows:

1. For $0 \leq \beta \leq \beta_c = N$, we obtain unique global minima for free energy at the origin with a zero magnetization.

2. For $\beta \geq \beta_c = N$, we have infinitely many global minima for the free energy which can be approximated graphically. Furthermore, the minima in this case are identical with non-zero magnetization.

Remarks on Curie-Weiss model ($N = 1$):

We can verify the expressions for magnetization and free energy for the Curie-Weiss model ($N = 1$) from expressions for N-vector model to verify that they matches with the Curie-Weiss model expressions.

1. The magnetization for the Curie-Weiss model is given by:

$$|m| = \frac{I_{\frac{1}{2}}(|m|)}{I_{\frac{1}{2}-1}(|m|)} = tanh(|m|).$$

2. The free energy φ for the N-vector model is:

$$\varphi_\beta(r) = \frac{1}{2N}\left(1 - \frac{\beta}{N}\right)r^2 + \frac{1}{N^2(N+2)}\left(-\frac{3}{4} + \frac{\beta}{N}\right)r^4 + O(r^6),$$

   which simplifies to $\varphi_1(r) = \frac{r^4}{12} + O(r^6)$ for the Curie-Weiss model at the critical point $\beta = N = 1$. This agrees with the density function for the Curie-Weiss model discussed in [VZ94, EN78b].

We can deduce from Proposition 3.1.1 the following Cramér-type LDP for the average spin $M_n := \frac{1}{n}S_n = \frac{1}{n}\sum_{i=1}^n \sigma_i$ in the noninteracting case $\beta = 0$ (or alternatively prove this Cramér theorem directly for random vectors on the hypersphere).

Corollary 3.1.4. *If $\{\sigma_i\}_{i=1}^n$ are i.i.d. uniform random points on $\mathbb{S}^{N-1} \subseteq \mathbb{R}^N$, then for $r = |x|$, the average spins M_n satisfy an LDP with rate function I:*

$$P_n\left(M_n \simeq x\right) \simeq e^{-nI(r)},$$

where $I(r) = \Phi_0(r)$ from Theorem 3.1.3.

The function $g(y) = \frac{I_1(y)}{I_0(y)}$ is strictly increasing on $(0, \infty)$, so that the equation above does uniquely define c as a function of $|x|$.

Similarly we have a Cramér-type result for the interacting case $\beta > 0$ as follows:

Proposition 3.1.5. *If $M_n = M_n(\sigma) := \frac{1}{n}\sum_{i=1}^n \sigma_i$ and $P_{n,\beta}$ is the Gibbs measure as defined above, then for a Borel set $\Gamma \in \mathbb{R}$,*

$$\begin{aligned} -\inf_{x\in\Gamma^\circ} I_\beta(x) &\leq \liminf_{n\to\infty} \frac{1}{n}\log P_{n,\beta}\left[\beta M_n \in \Gamma\right] \\ &\leq \limsup_{n\to\infty} \frac{1}{n}\log P_{n,\beta}\left[\beta M_n \in \Gamma\right] \leq -\inf_{x\in\overline{\Gamma}} I_\beta(x) \end{aligned}$$

where for $r = |x|$,

$$I_\beta(r) = \Phi_\beta(r) - \inf_{\nu \in M_1(\mathbb{S}^{N-1})} \Phi_\beta(r),$$

and $\Phi_\beta(r)$ is defined in Theorem 3.1.3.

For detailed theory of Cramér-type LDP and related rate function in statistical physics perspective, the reader is referred to the book by Ellis [Ell06].

In order to identify the explicit expression for the measure (similar to [KM13]), consider f to be a symmetric density function about the north pole which is decreasing away from the pole . We consider the measure ν_g with the density function $f(x_1, x_2, ..., x_N) = g(x_N)$ which is increasing in x_N. This gives:

$$\begin{aligned}
H(\nu_g \mid \mu) &= \int_{\mathbb{S}_{N-1}} f(x_1, x_2, ..., x_N) \log[f(x_1, x_2, ..., x_N)] dx_1 dx_2 ... dx_N \\
&= \frac{1}{A_N} \int_0^\pi \int_0^\pi ... \int_0^\pi \int_0^{2\pi} g(\cos(\theta_{N-1})) \log[g(\cos(\theta_{N-1}))] \\
&\qquad \prod_{k=2}^{N-1} sin^{k-1}(\theta_k) d\theta_1 d\theta_2 ... d\theta_{N-1} \\
&= \frac{A_{N-1}}{A_N} \int_0^\pi g(\cos(\theta_{N-1})) \log[g(\cos(\theta_{N-1}))] sin^{N-2}(\theta_{N-1}) d\theta_{N-1} \\
&= \frac{A_{N-1}}{A_N} \int_{-1}^1 g(x_N) \log[g(x_N)] \left(1 - x_N^2\right)^{\frac{N-3}{2}} dx_N.
\end{aligned}$$

Similarly, if e_N is the unit coordinate vector in direction x_N, then

$$\int_{\mathbb{S}^{N-1}} v d\nu_g(v) = e_N \frac{A_{N-1}}{A_N} \int_{-1}^1 x_N g(x_N) \left(1 - x_N^2\right)^{\frac{N-3}{2}} dx_N.$$

Now we are interested in minimizing the functional

$$\frac{A_{N-1}}{A_N} \int_{-1}^1 g(x_N) \log[g(x_N)] \left(1 - x_N^2\right)^{\frac{N-3}{2}} dx_N - \frac{\beta}{2} \left(\frac{A_{N-1}}{A_N} \int_{-1}^1 x_N g(x_N) \left(1 - x_N^2\right)^{\frac{N-3}{2}} dx_N \right)^2$$

under the following constraint: $g : [-1, 1] \to \mathbb{R}_+$ increasing and such that

$$\frac{A_{N-1}}{A_N} \int_{-1}^1 g(x_N) \left(1 - x_N^2\right)^{\frac{N-3}{2}} dx_N = 1.$$

We can also write the functional under consideration in terms of entropy as follows:

$$
\begin{aligned}
&\frac{A_{N-1}}{A_N}\int_{-1}^{1} g(x_N)\log[g(x_N)]\left(1-x_N^2\right)^{\frac{N-3}{2}}dx_N\\
&=\frac{A_{N-1}}{A_N}\int_{-1}^{1} g(x_N)\log\left[\frac{g(x_N)A_N}{A_{N-1}}\right]\left(1-x_N^2\right)^{\frac{N-3}{2}}dx_N+\log\left[\frac{A_{N-1}}{A_N}\right]\\
&=-\xi\left(\frac{gA_N}{A_{N-1}}\right)+\log\left[\frac{A_{N-1}}{A_N}\right],
\end{aligned}
$$

where $\xi(\phi)$ is the entropy of the density ϕ.

Now using constrained entropy optimization (Theorem 12.1.1 in [CT06]), we fix $c \in [0,1]$ and minimize the above quantity over the measures $\nu \in M_1(\mathbb{S}^{N-1})$ such that $\left|\int_{\mathbb{S}^{N-1}} x d\nu(x)\right| = c$.

Proposition 3.1.6. *Consider the set of functions $h : [-1,1] \to \mathbb{R}_+$ such that*

1. $\int_{-1}^{1} h(x_N)\left(1-x_N^2\right)^{\frac{N-3}{2}}dx_N = 1$, *and*
2. $\left|\int_{-1}^{1} x_N h(x_N)\left(1-x_N^2\right)^{\frac{N-3}{2}}dx_N\right| = c.$

Then in the set of functions satisfying these conditions, $h^*(x) = ae^{bx}$ uniquely minimizes the quantity

$$
\int_{-1}^{1} h(x_N)\log[h(x_N)]\left(1-x_N^2\right)^{\frac{N-3}{2}}dx_N.
$$

Now we will use the conditions (1) and (2) to find the values of the parameters a and b for the function h^*. The first condition leads us to the following two subcases: for N even,

$$
1=\int_{-1}^{1} ae^{bx_N}\left(1-x_N^2\right)^{\frac{N-3}{2}}dx_N=\frac{\left(\prod_{k=0}^{\frac{N}{2}-1}|2k-1|\right)a\pi I_{\frac{N}{2}-1}(b)}{b^{\frac{N}{2}-1}},
$$

which implies

$$
a=\frac{b^{\frac{N}{2}-1}}{\left(\prod_{k=0}^{\frac{N}{2}-1}|2k-1|\right)\pi I_{\frac{N}{2}-1}(b)}
$$

Now using second condition, For N even, we have

$$
c=\int_{-1}^{1} x_N h(x_N)\left(1-x_N^2\right)^{\frac{N-3}{2}}dx_N=\frac{\left(\prod_{k=0}^{\frac{N}{2}-1}|2k-1|\right)a\pi I_{\frac{N}{2}}(b)}{b^{\frac{N}{2}-1}}=\frac{I_{\frac{N}{2}}(b)}{I_{\frac{N}{2}-1}(b)}.
$$

Let $g^* = \left(\frac{A_{N-1}}{A_N}\right)h^*$; and $c \in [0,1]$ with g^* increasing corresponds to considering all $b \in [0,\infty)$. Now we have to minimize for N even,

$$
\begin{aligned}
&\int_{-1}^{1} g^*(x_N)\log\left[\frac{A_N}{A_{N-1}}g^*(x_N)\right]\left(1-x_N^2\right)^{\frac{N-3}{2}}dx_N-\frac{\beta}{2}\left(\int_{-1}^{1}x_N g^*(x_N)\left(1-x_N^2\right)^{\frac{N-3}{2}}dx_N\right)^2\\
&=b\frac{I_{\frac{N}{2}}(b)}{I_{\frac{N}{2}-1}(b)}+\log\left[\frac{A_N}{A_{N-1}}\frac{b^{\frac{N}{2}-1}}{\left(\prod_{k=0}^{\frac{N}{2}-1}|2k-1|\right)\pi I_{\frac{N}{2}-1}(b)}\right]-\frac{\beta}{2}\left(\frac{I_{\frac{N}{2}}(b)}{I_{\frac{N}{2}-1}(b)}\right)^2 \qquad (3.4)\\
&=:\Phi_\beta(b)
\end{aligned}
$$

over all $b\in[0,\infty)$. Using a similar approach we can derive the expression for $\Phi_\beta(b)$, for the other subcase N odd, this comes out to be, for all N,

$$
\Phi_\beta(b)=b\frac{I_{\frac{N}{2}}(b)}{I_{\frac{N}{2}-1}(b)}+\log\left[\frac{A_N}{A_{N-1}}\frac{b^{\frac{N}{2}-1}}{B_N\ \pi\ I_{\frac{N}{2}-1}(b)}\right]-\frac{\beta}{2}\left(\frac{I_{\frac{N}{2}}(b)}{I_{\frac{N}{2}-1}(b)}\right)^2,
$$

where

$$
B_N=\begin{cases}\prod_{k=0}^{\frac{N}{2}-1}|2k-1|, & \text{if } N \text{ even},\\[2ex] \frac{2^{\frac{N}{2}-1}\Gamma\left(\frac{N-1}{2}\right)}{\sqrt{\pi}}, & \text{if } N \text{ odd}.\end{cases}
$$

This is now one-dimensional problem which is studied in Lemma 5.1.1 in chapter 5. We deduce that $\beta_c=N$ is the critical inverse temperature. Also for $\beta<N$ we have the uniform distribution as the only macrocanonical state whereas for $\beta>N$ we have a family (parametrized by the circle) of distributions with a preferred direction (and converging to a family of points masses, each concentrated on a perfectly preferred direction as $\beta\to\infty$). We can state the following theorem using our calculations from Lemma 5.1.1:

Theorem 3.1.7. *1. For $\beta\le N$, the expression* (3.4) *is minimized for $b=0$, then the corresponding $a=1$, so that the minimizing function $h^*=1$ and hence the canonical macrostates in the subcritical case are uniform: $\mathcal{E}_\beta=\{\mu\}$.*

2. In the supercritical case, $\beta>N$, the minimizing b for the expression (3.4) *is the unique strictly positive solution to*

$$
b=\beta\left(\frac{I_{\frac{N}{2}}(b)}{I_{\frac{N}{2}-1}(b)}\right),
$$

which moreover has limit $\lim_{\beta\downarrow\beta_c}b=0$. The macrostates $\mathcal{E}_\beta$ are given by $\{\nu_x\}_{x\in\mathbb{S}^{N-1}}$, where ν_x is the probability measure with density which is symmetric about the pole at x, with density $g_x:[-1,1]\to\mathbb{R}$ in the x-direction given by $\left(\frac{A_{N-1}}{A_N}\right)ae^{bx}$ with b as above.

3.2 Limit Theorems for the Total Spin

We study the total spin in the subcritical, critical, and supercritical regimes, proving central and non-central limit theorems for the total spin, holding N, the dimension of the spin, fixed. In this section we state these limit theorems for each regime and give the proofs of each one of these in next few sections.

In the subcritical regime $0 \le \beta < N$, the spins are weakly correlated and hence can be treated similar to the independent case $\beta = 0$. The average magnetization of the system is very small and goes to zero with increasing number of spins $n \to \infty$ for this high temperature regime. In this regime we have the following multivariate central limit theorem, and in particular, the macrostate is the uniform measure on the hypersphere.

Theorem 3.2.1. *In the subcritical regime $\beta < N$, the random variable W_n is defined as follows: $W_n = \sqrt{\frac{N-\beta}{n}} \sum_{i=1}^{n} \sigma_i$ Then*

$$\sup_{g:L(g),M(g)\le 1} |\,\mathbb{E}\, g(W_n) - \mathbb{E}\, g(Z)| \le \frac{c_\beta}{\sqrt{n}}$$

where c_β is a constant depending only on β, $L(g)$ is the Lipschitz constant of g, $M(g)$ is the maximum operator norm of the Hessian of g, and Z is a standard Gaussian random vector in $\mathbb{R}^N$.

Remark: Our rate of convergence for Theorem 3.2.1 is sharper than [KM13], which had a factor of $\log(n)$ in the numerator, based on an argument of Leslie Ross [Ros17]. The supremum in Theorem 3.2.1 is a metric for the topology of weak-* convergence and convergence in mean on the space $M_1(\mathbb{S}^{N-1})$ of probability measures on the hypersphere.

In the supercritical regime $\beta > N$, the spins align to some extent: For smaller values of $\beta > N$, the spins show a slight preference for a particular (random) direction, whereas for large β, the spins align strongly. Consider a small interval Γ containing b, now using the fact that $\inf_{x\in\overline{\Gamma}} I_\beta(x) = b$ and Proposition 3.1.5, we conclude that $|S_n|$ has a high probability of being close to bn/β. Here all points on the hypersphere of radius bn/β will have equal probability due to symmetry. Using an argument similar to [KM13], we consider the fluctuations of squared-length of total spin, i.e., we consider the following random variable:

$$W_n := \sqrt{n}\left[\frac{\beta^2}{n^2 b^2}\left|\sum_{j=1}^{n} \sigma_j\right|^2 - 1\right]. \tag{3.5}$$

In Section 3.4, we prove that this W_n satisfies the following central limit theorem.

Theorem 3.2.2. *If W_n is as defined in* (3.5) *and b is the solution of $b - \beta f(b) = 0$, where*

$$f(b) = \frac{I_{\frac{N}{2}}(b)}{I_{\frac{N}{2}-1}(b)},$$

then there is a constant c_β depending on $\beta > N$ only, such that if Z is a centered normal random variable with variance

$$Var(\sigma) = \frac{4\beta^2}{(1-\beta f'(b))\,b^2}\left[1 - \frac{N-1}{b}\frac{I_{\frac{N}{2}}(b)}{I_{\frac{N}{2}-1}(b)} - \left(\frac{I_{\frac{N}{2}}(b)}{I_{\frac{N}{2}-1}(b)}\right)^2\right]$$

then

$$d_{BL}(W_n, Z) \leq c_\beta \left(\frac{\log(n)}{n}\right)^{1/4}.$$

Here the bounded Lipschitz distance $d_{BL}(X,Y)$ between random variables X and Y is:

$$d_{BL}(X,Y) := \sup\left\{\left|\mathbb{E}\,h(X) - \mathbb{E}\,h(Y)\right| : \|h\|_\infty \leq 1, L(h) \leq 1\right\},$$

where $\|\cdot\|_\infty$ is the supremum norm and $L(\cdot)$ is the Lipschitz constant as before.

We can obtain the complete asymptotic behavior of the total spin without using conditioning (as in e.g., [EN78a]) by using instead the rotational invariance of the total spin, a strategy adapted from [KM13].

In Section 3.5, we prove the following nonnormal limit theorem for the random variable defined by

$$W_n := \frac{c_N |S_n|^2}{n^{\frac{3}{2}}}.$$

at the critical temperature $\beta = N$. Because of symmetry of the total spin this leads us to the limiting picture in the critical case. The critical limiting density function p (defined below) is obtained using Stein's method similar to [CS11, KM13].

Theorem 3.2.3. *If we consider the critical temperature $\beta = N$, and W_n as defined by* (2.7) *, and if X is the random variable with the density*

$$p(t) = \begin{cases} \frac{1}{z} t^{\frac{N-2}{2}} e^{-\frac{1}{4N^2(N+2)}t^2} & t \geq 0; \\ 0 & t < 0, \end{cases},$$

where z is normalizing constant and c_N is such that $\mathbb{E}\,W_n = 1$, then there exists a universal constant C such *that*

$$\sup_{\substack{\|h\|_\infty \le 1,\, \|h'\|_\infty \le 1 \\ \|h''\|_\infty \le 1}} \left| \mathbb{E}\, h(W_n) - \mathbb{E}\, h(X) \right| \le \frac{C \log(n)}{\sqrt{n}}.$$

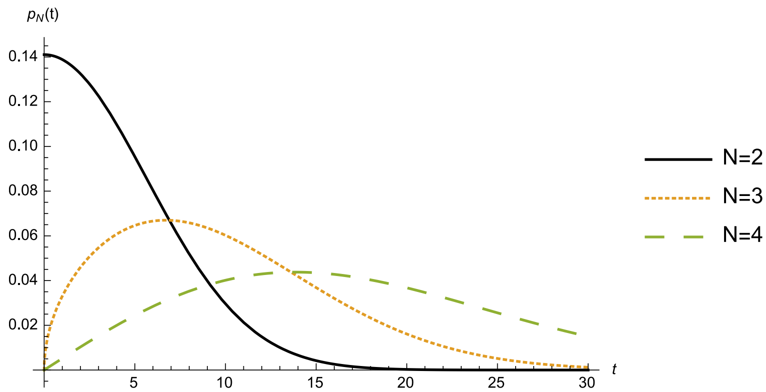

Figure 3.4: Mean-field critical density functions p_N for $2 \le N \le 4$ and $t \ge 0$. For the XY model $p_2(t) = \frac{e^{-t^2/64}}{4\sqrt{\pi}}$, for the Heisenberg model $p_3(t) = \frac{\sqrt{t}e^{-t^2/180}}{5^{3/4}\sqrt{54}\Gamma[3/4]}$, and for the Toy model of the Higgs sector, $p_4(t) = \frac{te^{-t^2/384}}{192}$.

Remarks on density function:

We would like to note that we also have made correction to the critical distribution function given in [KM13]. The mistake in [KM13] was a wrong heuristic used to predict the distribution $p(t)$, and here in our present manuscript, we use a different, more reliable method to derive $p(t)$, whose main idea is to use the Taylor series expansion of the free energy as discussed by Ellis and Newman [EN78b]. So we have expressed our free energy function in Taylor series to find the values of $m, k, \lambda(0)$ and then used the formula given in [EN78b].

3.3 The Subcritical Phase

This section has the proof of Theorem 3.2.1, the limit theorem for S_n in the disordered phase. We start by calculating the variance of the total spin $S_n := \sum_{i=1}^n \sigma_i$. Since the density of the Gibbs measure is symmetric and in particular rotationally invariant, each of the spins σ_i has a uniform marginal distribution, and $\mathbb{E}\langle\, \sigma_i, \sigma_i \rangle\ = 1$ for each i and $\mathbb{E}\langle\, \sigma_i, \sigma_j \rangle$ is the same for every pair $i \neq j$. In particular, $\mathbb{E}[\sigma_i] = 0$ for each i, and thus the expected total spin is indeed zero.

Following [KM13], the density of σ_1 with respect to uniform measure on $\mathbb{S}^{N-1}$, conditional on $\{\sigma_j\}_{j \neq 1}$, is

$$Z_1^{-1} \exp\left[\frac{\beta}{n} \sum_{j \neq 1} \langle\, \theta, \sigma_j \rangle\right],$$

where $Z_1 = \int_{\mathbb{S}^{N-1}} \exp\left[\frac{\beta}{n}\sum_{j\neq 1}\langle\, \theta, \sigma_j\rangle\right] d\mu(\theta)$ is the normalization factor. If $i \in \{1,\dots,n\}$ is fixed, then call $\sigma^{(i)} := \sum_{j\neq i}\sigma_j$. We use hyperspherical coordinates,

$$d\mu(\theta) = \frac{J_N}{A_N} d\theta_1 \dots d\theta_{N-1},$$

where

$$J_N = (-1)^{N-1}\prod_{k=2}^{N-1}\sin^{k-1}(\theta_k).$$

Here $A_N = 2\pi^{\frac{N}{2}}/\Gamma\left(\frac{N}{2}\right)$, and we also use the notation $\kappa = \frac{\beta|\sigma^{(1)}|}{n}$. Therefore the normalization factor is

$$\begin{aligned}
Z_1 &= \frac{1}{A_N}\int_0^{\pi}\dots\int_0^{\pi}\int_0^{2\pi} e^{\kappa\cos(\theta_{N-1})} J_N d\theta_1 ... d\theta_{N-2} d\theta_{N-1} \\
&= \frac{A_{N-1}}{A_N}\int_0^{\pi} e^{\kappa\cos(\theta_{N-1})}\sin^{N-2}(\theta_{N-1})d\theta_{N-1} \\
&= \frac{A_{N-1}}{A_N}\int_{-1}^{1} e^{\kappa u}\left(1-u^2\right)^{\frac{N-3}{2}} du. \\
&= \frac{\left(\prod_{j=0}^{\frac{N}{2}-1}(2j+1)\right) a\pi I_{\frac{N}{2}-1}(\kappa)}{\kappa^{\frac{N}{2}-1}}
\end{aligned}$$

The conditional expectation can be calculated using the conditional density as follows:

$$\begin{aligned}
\mathbb{E}\left[\sigma_1 \middle| \{\sigma_j\}_{j\neq 1}\right] &= \frac{1}{Z_1}\int_{\mathbb{S}^{N-1}} \theta \exp\left[\frac{\beta}{n}\sum_{j\neq 1}\langle\, \theta, \sigma_j\rangle\right] d\mu(\theta) \\
&= \frac{I_{\frac{N}{2}}(\kappa)}{I_{\frac{N}{2}-1}(\kappa)}\cdot\frac{\sigma^{(1)}}{|\sigma^{(1)}|}.
\end{aligned}$$

Here again $I_{\frac{N}{2}}$ is the modified Bessel function of the first kind. A series expansion about zero gives $\frac{I_{\frac{N}{2}}(\kappa)}{I_{\frac{N}{2}-1}(\kappa)} \approx \frac{\kappa}{N}$ for small κ, hence for $\beta < N$, we have

$$\mathbb{E}\left[\sigma_1 \middle| \{\sigma_j\}_{j\neq 1}\right] \approx \frac{\kappa}{N} = \frac{\beta\sigma^{(1)}}{Nn} = \frac{\beta}{Nn}\sum_{i\neq 1}\sigma_i,$$

and taking an inner product with σ_2, taking expectation, and using symmetry we obtain:

$$\mathbb{E}\left[\langle\, \sigma_1, \sigma_2\rangle | \{\sigma_j\}_{j\neq 1}\right] \approx \frac{\beta}{Nn}\sum_{i\neq 1}\langle\, \sigma_i, \sigma_2\rangle.$$

$$\mathbb{E}\langle\, \sigma_1,\sigma_2\rangle = \mathbb{E}\left[\mathbb{E}\left[\langle\, \sigma_1,\sigma_2\rangle|\{\sigma_j\}_{j\neq 1}\right]\right] \approx \frac{\beta}{Nn}\mathbb{E}\langle\, \sum_{j\neq 1}\sigma_j, \sigma_2\rangle = \frac{\beta}{Nn}\left[1 + (n-2)\,\mathbb{E}\langle\, \sigma_1,\sigma_2\rangle\right],$$

and thus

$$\mathbb{E}\langle \sigma_1, \sigma_2\rangle \approx \frac{\beta}{Nn - \beta(n-2)} \approx \frac{\beta}{n(N-\beta)}. \tag{3.6}$$

Finally,

$$\mathbb{E}\,|S_n|^2 = n\,\mathbb{E}\,|\sigma_1|^2 + n(n-1)\,\mathbb{E}\langle \sigma_1, \sigma_2\rangle \approx \frac{2n}{N-\beta}.$$

Theorem 3.2.1 is an application of an abstract normal approximation theorem from [Mec09], a version of Stein's method of exchangeable pairs [Ste86]. The specific version used on the analogous mean-field Heisenberg model is Theorem 14 in [KM13].

We need to construct an exchangeable pair (W_n, W_n') for applying these theorems [KM13, Mec09]. Using Gibbs sampling, we start with a configuration σ and construct a new configuration σ' that differs at only one site by picking I uniformly at random in $\{1, \ldots, n\}$ and replacing the original spin σ_I by the new spin σ_I'. The total spin of the original configuration is $W_n = \sqrt{\frac{N-\beta}{n}} \sum_{i=1}^{n} \sigma_i$ and the total spin of the new configuration is

$$W_n' = W_n(\sigma') = W_n - \sqrt{\frac{N-\beta}{n}}\sigma_I + \sqrt{\frac{N-\beta}{n}}\sigma_I'.$$

The lemma below gives expressions for the quantities R, R', and Λ appearing in the cited theorems [KM13, Mec09].

Lemma 3.3.1. *If the exchangeable pair (W_n, W_n') is obtained using the Gibbs sampling construction above, and $f(\kappa) = \frac{I_{\frac{N}{2}}(\kappa)}{I_{\frac{N}{2}-1}(\kappa)}$ and $\Lambda = \left(\frac{1-\frac{\beta}{N}}{n}\right) Id$, then*

1. $\mathbb{E}\left[W_n' - W_n \middle| \sigma\right] = -\Lambda W_n + R$, *where*

$$R = -\frac{\beta}{Nn^2}W_n - \frac{\beta^3}{N^2(N+2)n^2}\left(W_n - \frac{W_n}{n}\right) + \frac{a}{n^{3/2}}\sum_{i=1}^{n}\left[f(\kappa) - \frac{\kappa}{N} + \frac{\kappa^3}{N^2(N+2)}\right]\frac{\sigma^{(i)}}{|\sigma^{(i)}|}.$$

2. $\mathbb{E}\left[(W_n' - W_n)(W_n' - W_n)^T \middle| \sigma\right] = 2\Lambda + R'$, *with*

$$\begin{aligned} R' = {} & \frac{a^2}{Nn}\left[\frac{1}{n}\sum_{i=1}^{n} N\sigma_i\sigma_i^T - Id\right] - \left[\frac{2\beta}{Nn^2}W_nW_n^T - \frac{2a^2\beta}{Nn^3}\sum_{i=1}^{n}\sigma_i\sigma_i^T\right] \\ & + \frac{a^2}{n^2}\sum_{i=1}^{n}\left\{\left[\frac{I_{\frac{N}{2}}(\kappa) + \kappa I_{\frac{N}{2}+1}(\kappa)}{\kappa I_{\frac{N}{2}-1}(\kappa)} - \frac{1}{N}\right]P_i + \left[\frac{I_{\frac{N}{2}}(\kappa)}{cI_{\frac{N}{2}-1}(\kappa)} - \frac{1}{N}\right]P_i^{\perp}\right. \\ & \left. - \left[\frac{I_{\frac{N}{2}}(\kappa)}{I_{\frac{N}{2}-1}(\kappa)} - \frac{\kappa}{N}\right](r_i\sigma_i^T + \sigma_i r_i^T)\right\}. \end{aligned}$$

Now we will give the bounds for R and R' calculated as above.

Lemma 3.3.2. *For the exchangable pair (W_n, W_n') which is constructed using Gibbs sampling and R, R' as in the previous lemma, there is a constant c_β such that*

1. $\mathbb{E}\,|R| \le \frac{c_\beta(N)}{n^{3/2}}$;
2. $\mathbb{E}\,\|R'\|_{HS} \le \frac{c_\beta(N)}{n^{3/2}}$;
3. $\mathbb{E}\,|W_n' - W_n|^3 \le \frac{c_\beta(N)}{n^{3/2}}$.

Theorem 3.2.1 follows from the abstract normal approximation theorem in [Mec09] and Lemmas 3.3.1 and 3.3.2. The detailed proofs for Lemma 3.3.1 and 3.3.2 are given in the chapter 4.

3.4 The Supercritical Phase

For proving theorem 3.2.2, we will use a version of Stein's abstract normal approximation theorem [Ste86] (p.35). The formulation given below is a univariate analog of abstract normal approximation theorem from [Mec09].

Consider the random variable $W_n = \sqrt{n}\left[\frac{\beta^2}{n^2 b^2}\left|\sum_{j=1}^n \sigma_j\right|^2 - 1\right]$, for supercritical case as explained in section 3. Now construct an exchangable pair (W_n, W_n') using Glauber dynamics in order to apply Stein's abstract normal approximation theorem to W_n. We will describe the lemma which contains the bounds needed to obtain Theorem 3.2.2 from Stein abstract theorem. This will lead us to proof of Theorem 3.2.2.

Lemma 3.4.1. *Define $f(r) = \frac{I_{\frac{N}{2}}(r)}{I_{\frac{N}{2}-1}(r)}$, and let b be the positive solution of*

$$b - \beta f(b) = 0.$$

Then for the exchangable pair (W_n, W_n') as constructed above,

1. *For $\lambda = \frac{1-\beta f'(b)}{n}$,*

$$\mathbb{E}\left[W_n' - W_n \middle| \sigma\right] = -\lambda W_n + R \qquad \text{and} \qquad \mathbb{E}\,|R| \le \frac{c_\beta(N)\log(n)}{n^{3/2}};$$

2. *For $Var(\sigma) = \frac{4\beta^2}{(1-\beta f'(b))b^2}\left[1 - \frac{N-1}{N}\frac{\left(I_{\frac{N}{2}-1}(b)+I_{\frac{N}{2}+1}(b)\right)}{I_{\frac{N}{2}-1}(b)} - \left(\frac{I_{\frac{N}{2}}(b)}{I_{\frac{N}{2}-1}(b)}\right)^2\right]$,*

$$\mathbb{E}\left|\sigma^2 - \frac{1}{2\lambda}\,\mathbb{E}\left[(W_n' - W_n)^2\middle|\sigma\right]\right| \le \frac{c_\beta(N)(\log(n))^{1/4}}{n^{1/4}};$$

3. $\mathbb{E}\,|W_n' - W_n|^3 \leq \frac{c_\beta}{n^{3/2}}$.

Proof. First of all λ and σ defined above are always strictly positive. For par (a), first we found the bounds for $f(x)$ and then using LDP for $|S_n|$ along with taylor series expansion we can deduce the result. Consider,

$$\begin{aligned}
\mathbb{E}\left[W_n' - W_n \middle| \sigma\right] &= -\frac{\beta^2}{n^{5/2}b^2}\sum_{i=1}^{n}\left[2\sum_{k\neq i}\langle\ \sigma_i, \sigma_k\rangle - \mathbb{E}\left[2\sum_{k\neq i}\langle\ \sigma_i, \sigma_k\rangle \middle| \{\sigma_j\}_{j\neq i}\right]\right] \\
&= -\frac{2\beta^2}{n^{5/2}b^2}\left(\left|\sum_{i=1}^{n}\sigma_i\right|^2 - n\right) + \frac{2\beta^2}{n^{5/2}b^2}\sum_{i=1}^{n} f\left(\frac{\beta|\sigma^{(i)}|}{n}\right)|\sigma^{(i)}| \\
&= -\frac{2}{n}W_n - \frac{2}{\sqrt{n}} + \frac{2\beta^2}{n^{3/2}b^2} + \frac{2\beta^2}{n^{5/2}b^2}\sum_{i=1}^{n} f\left(\frac{\beta|\sigma^{(i)}|}{n}\right)|\sigma^{(i)}|.
\end{aligned}$$

For $f(r)$, using the same approach as in section 5 of [KM13], we have

$$\begin{aligned}
&\frac{2\beta^2}{n^{5/2}b^2}\sum_{i=1}^{n}\mathbb{E}\left| f\left(\frac{\beta|\sigma^{(i)}|}{n}\right)|\sigma^{(i)}| - f\left(\frac{\beta|S_n|}{n}\right)|S_n|\right| \\
&\qquad \leq \frac{2\|f'\|_\infty\beta^3 + 2\|f\|_\infty\beta^2}{n^{3/2}b^2};
\end{aligned}$$

that is,

$$\mathbb{E}\left[W_n' - W_n \middle| \sigma\right] = -\frac{2}{n}W_n - \frac{2}{\sqrt{n}} + \frac{2\beta^2}{n^{3/2}b^2} f\left(\frac{\beta|S_n|}{n}\right)|S_n| + R_1, \tag{3.7}$$

where $\mathbb{E}\,|R_1| \leq \frac{c_\beta}{n^{3/2}}$.

Now we will use Taylor expansion to approximate $f\left(\frac{\beta|S_n|}{n}\right)$ using the LDP for $|S_n|$ (Proposition 3.1.5). For $r = |x|$, we obtain

$$\limsup_{n\to\infty}\frac{1}{n}P_{n,\beta}\left[\left|\frac{|S_n|}{n} - \frac{b}{\beta}\right| \geq \epsilon\right] \leq -\inf_{\left|r-\frac{b}{\beta}\right|\geq\epsilon} I_\beta(r),$$

where

$$I_\beta(r) = \Phi_\beta(y) - \varphi(\beta),$$

with

$$\Phi_\beta(y) = r\ y + \log\left[\frac{A_N}{A_{N-1}}\frac{y^{\frac{N}{2}-1}}{B_N\pi I_{\frac{N}{2}-1}(y)}\right] - \frac{\beta}{2}r^2,$$

here B_N is defined in (3.3) and y is calculated from

$$\frac{I_{\frac{N}{2}}(y)}{I_{\frac{N}{2}-1}(y)} = r,$$

and

$$\varphi(\beta) = \inf_{x \geq 0} \Phi_\beta(y)$$

Note that $\Phi_\beta(r)$ is decreasing on $\left[0, \frac{b}{\beta}\right]$ and increasing on $\left[\frac{b}{\beta}, \infty\right)$. Also at $y = b$, $r = \frac{b}{\beta}$ is the unique minimizing set for Φ_β. That is, for $I_\beta(y(t)) = I_\beta(f^{-1}(t))$, we have

$$\inf_{\left|r - \frac{b}{\beta}\right| \geq \epsilon} I_\beta(r) = \min\left\{ I_\beta\left(y\left(\frac{b}{\beta} - \epsilon\right)\right), I_\beta\left(y\left(\frac{b}{\beta} + \epsilon\right)\right)\right\}.$$

This implies $\Phi'_\beta(b) = 0$. Furthermore, $\Phi''_\beta(b) > 0$, which implies that there is a constant $C_\beta(N)$ such that

$$\inf_{\left|r - \frac{b}{\beta}\right| \geq \epsilon} I_\beta(r) \geq C_\beta(N)\epsilon^2,$$

which leads to

$$P_{n,\beta}\left[\left|\frac{|S_n|}{n} - \frac{b}{\beta}\right| \geq \epsilon\right] \leq e^{-C_\beta(N)n\epsilon^2}.$$

Now the approach similar to section 5 of [KM13], where we use $|S_n| \leq n$ and $f(r) = \frac{r}{\beta}$ in equation (3.7) leads to

$$\mathbb{E}\left[W'_n - W_n \middle| \sigma\right] = -\frac{1 - \beta f'(b)}{n} W_n + R,$$

where again $\mathbb{E}|R| \leq \frac{c_\beta(N)\log(n)}{n^{3/2}}$. This completes the proof of part (a).

For part (b), we will show the positivity of σ^2 and then we will use the asymptotics expanion of $|\sigma_i|$ and $\langle\ \sigma_i, \sigma^{(i)}\rangle$ in order to write the bound for second moment. Observe that by definition,

$$\begin{aligned}
\mathbb{E}\left[(W'_n - W_n)^2 \middle| \sigma\right] &= \frac{\beta^4}{n^4 b^4} \sum_{i=1}^n \mathbb{E}\left[\left(2\sum_{j\neq i}\langle\ \sigma_i^* - \sigma_i, \sigma_j\rangle\right)^2 \middle| \sigma, I = i\right] \\
&= \frac{4\beta^4}{n^4 b^4} \sum_{i=1}^n \sum_{j,k\neq i} \mathbb{E}\left[\sigma_j^T(\sigma_i^* - \sigma_i)(\sigma_i^* - \sigma_i)^T\sigma_k \middle| \sigma, I = i\right].
\end{aligned} \tag{3.8}$$

Notice that here σ_i^* is coming from the definition of the exchangabale pair $(W_n,\ W'_n)$. From the subcritical case calculations we have

$$\begin{aligned}
&\mathbb{E}\left[(\sigma_i^* - \sigma_i)(\sigma_i^* - \sigma_i)^T \middle| \sigma, I = i\right] \\
&\quad = \left\{\left[\frac{I_{\frac{N}{2}}(\kappa) + \kappa I_{\frac{N}{2}+1}(\kappa)}{\kappa I_{\frac{N}{2}-1}(\kappa)}\right] P_i + \left[\frac{I_{\frac{N}{2}}(\kappa)}{\kappa I_{\frac{N}{2}-1}(\kappa)}\right] P_i^\perp - \left[\frac{I_{\frac{N}{2}}(\kappa)}{I_{\frac{N}{2}-1}(\kappa)}\right](r_i\sigma_i^T + \sigma_i r_i^T) + \sigma_i\sigma_i^T\right\}.
\end{aligned}$$

where $\kappa = \frac{\beta|\sigma^{(i)}|}{n}$, $r_i = \frac{\sigma^{(i)}}{|\sigma^{(i)}|}$, and P_i is orthogonal projection onto r_i. Since W_n is defined differently for supercritical case, we will later substitute modification of above expression into (3.8).

In order to make sure that $\sigma^2 > 0$, we rewrite the first term of the last expression as

$$\frac{I_{\frac{N}{2}}(\kappa) + \kappa I_{\frac{N}{2}+1}(\kappa)}{\kappa I_{\frac{N}{2}-1}(\kappa)} = 1 - \frac{\kappa I_{\frac{N}{2}-1}(\kappa) - I_{\frac{N}{2}}(\kappa) - \kappa I_{\frac{N}{2}+1}(\kappa)}{\kappa I_{\frac{N}{2}-1}(\kappa)}.$$

Using series expansions of each term we have

$$\begin{aligned}\mathbb{E}\left[(\sigma_i^* - \sigma_i)(\sigma_i^* - \sigma_i)^T \middle| \sigma, I = i\right] &= \left(1 - \frac{N-1}{\beta}\right) P_i + \frac{1}{\beta} P_i^{\perp} - \frac{b}{\beta}(r_i \sigma_i^T + \sigma_i r_i^T) + \sigma_i \sigma_i^T + R_i' \\ &= \frac{1}{\beta} Id + \left(1 - \frac{N}{\beta}\right) P_i - \frac{b}{\beta}(r_i \sigma_i^T + \sigma_i r_i^T) + \sigma_i \sigma_i^T + R_i',\end{aligned} \tag{3.9}$$

where

$$R_i' = \left\{ \left[\frac{I_{\frac{N}{2}}(\kappa) + \kappa I_{\frac{N}{2}+1}(\kappa)}{\kappa I_{\frac{N}{2}-1}(\kappa)} - \frac{1}{\beta}\right] P_i + \left[\frac{I_{\frac{N}{2}}(\kappa)}{\kappa I_{\frac{N}{2}-1}(\kappa)} - \frac{1}{\beta}\right] P_i^{\perp} - \left[\frac{I_{\frac{N}{2}}(\kappa)}{I_{\frac{N}{2}-1}(\kappa)} - \frac{b}{\beta}\right] (r_i \sigma_i^T + \sigma_i r_i^T) \right\}.$$

Using the main term of (3.9) into (3.8) yields

$$\begin{aligned}&\mathbb{E}\left[(\sigma_i^* - \sigma_i)(\sigma_i^* - \sigma_i)^T \middle| \sigma, I = i\right] \\ &= \frac{4\beta^4}{n^4 b^4} \sum_i \sum_{j,k \neq i} \left[\frac{1}{\beta}\langle\, \sigma_j, \sigma_k\rangle + \left(1 - \frac{N}{\beta}\right) \sigma_j^T P_i \sigma_k - \frac{b}{\beta}(\sigma_j^T r_i \sigma_i^T \sigma_k + \sigma_j^T \sigma_i r_i^T \sigma_k) + \sigma_j^T \sigma_i \sigma_i^T \sigma_k\right].\end{aligned}$$

Following the calculations from section 5 of [KM13], we get the following simplified form

$$\begin{aligned}\mathbb{E}\left[(W_n' - W_n)^2 \middle| \sigma\right] = \frac{4\beta^4}{n^4 b^4} \sum_i &\left[\left(1 - \frac{N-1}{\beta}\right) |\sigma^{(i)}|^2 - \frac{2b}{\beta} |\sigma^{(i)}| \langle\, \sigma_j, \sigma^{(i)}\rangle + \langle\, \sigma_i, \sigma^{(i)}\rangle^2\right] \\ &+ \frac{4\beta^4}{n^4 b^4} \sum_i \sum_{j,k \neq i} \sigma_j^T R_i' \sigma_k.\end{aligned}$$

Now we have to find the deterministic constant which will be used to approximate the above final expression. Since $|\sigma^{(i)}| \approx \frac{b(n-1)}{\beta}$ for each i, and $\langle\, \theta, \sigma^{(i)}\rangle = \langle\, \sigma_i, S_n\rangle - 1$, this implies that $\langle\, \sigma_i, \sigma^{(i)}\rangle \approx \frac{|S_n|^2}{n} - 1 \approx \frac{nb^2}{\beta^2} - 1$. We also have to rewrite the last expression in a deterministic way such that we can represent it in the form $2\lambda Var(\sigma)$ plus a mean zero term. It is important to note that this will help us to find the value of $Var(\sigma)$.

Therefore we rewrite the above expression as:

$$\begin{aligned}
&\mathbb{E}\left[(W_n' - W_n)^2 \middle| \sigma\right] \\
&= \frac{4\beta^4}{n^3 b^4}\left[2\left(1-\frac{N-1}{\beta}\right)\frac{(n-1)^2 b^2}{\beta^2} - \frac{2n^2 b^4}{\beta^4}\right] + \frac{4\beta^4}{n^4 b^4}\sum_i \left(1-\frac{N-1}{\beta}\right)\left(|\sigma^{(i)}|^2 - \frac{(n-1)^2 b^2}{\beta^2}\right) \\
&\quad + \frac{4\beta^4}{n^4 b^4}\sum_i \left[-\frac{2b}{\beta}\left(|\sigma^{(i)}|\langle\, \sigma_i, \sigma^{(i)}\rangle - \frac{n^3 b^3}{\beta^4}\right) + \langle\, \sigma_i, \sigma^{(i)}\rangle^2 - \left(1-\frac{N-1}{\beta}\right)\frac{(n-1)^2 b^2}{\beta^2}\right] \\
&\quad + \frac{4\beta^4}{n^4 b^4}\sum_i \sum_{j,k\neq i} \sigma_j^T R_i' \sigma_k,
\end{aligned} \tag{3.10}$$

and similar to [KM13] we define σ such that we can express the leading order term of (3.10) as:

$$\frac{4\beta^4}{n^3 b^4}\left[2\left(1-\frac{N-1}{\beta}\right)\frac{(n-1)^2 b^2}{\beta^2} - \frac{2n^2 b^4}{\beta^4}\right] = 2\lambda Var(\sigma), \tag{3.11}$$

where $\lambda = \frac{1-\beta f'(b)}{n}$ defined as in part(a). Using the fact that $f(r) = \frac{I_{\frac{N}{2}}(r)}{I_{\frac{N}{2}-1}(r)} \approx \frac{r}{\beta}$, equation (3.11) simplifies to:

$$2\lambda Var(\sigma) = \frac{2n^2 b^2}{\beta^2}\left[1 - \frac{(N-1)\left(\frac{I_{\frac{N}{2}}(b)}{I_{\frac{N}{2}-1}(b)}\right)}{b} - \left(\frac{I_{\frac{N}{2}}(b)}{I_{\frac{N}{2}-1}(b)}\right)^2\right].$$

Let $\nu = \frac{N}{2} - 1$, then the positivity of the last expression was proved while deriving the improved bounds on the ratio of Bessel functions by Amos [Amo74a].

This yield a strictly positive value of $Var(\sigma)$ which depends only on β and N and is independent of n. Now for applying Theorem from [Ste86] (p.35) we need to estimate the expected absolute value of each of the terms above, which is straightforward and similar to [KM13], calculation for corresponding section [5] except all the $1 - \frac{2}{\beta}$ are replaced with $1 - \frac{N-1}{\beta}$ which only changes the vaue of c_β.

Finally, part (c) is trivial and similar to [KM13] section 5, with a different variance $Var(\sigma)$ coming from the corresponding hypersphere. $\square$

3.5 The Critical Temperature

Theorem 3.5.1. *Consider an exchangable pair of positive random variables (W, W'). Assume that there exists a σ-field $\mathcal{F} \supseteq \sigma(W)$, such that*

$$\mathbb{E}\left[W' - W \middle| \mathcal{F}\right] = Nk\big(1 - cW^2\big) + R$$

and

$$\mathbb{E}\left[(W'-W)^2\big|\mathcal{F}\right] = kW + R'.$$

where R and R' are $\mathcal{F}$-measurable random variables and $k > 0$ deterministic. Now consider a random variable X with density function

$$p(t) = \begin{cases} \frac{1}{z} t^{\frac{N-2}{2}} e^{-\frac{t^2}{4N^2(N+2)}} & t \geq 0; \\ 0 & t < 0, \end{cases}$$

where z is normalizing constant. Then there are constants C_1, C_2, C_3 such that for all $h \in C^2(\mathbb{R})$,

$$\begin{aligned} \big|\,\mathbb{E}\,h(W) - \mathbb{E}\,h(X)\big| \leq \frac{C_1\|h\|_\infty}{k}\,\mathbb{E}\,|R| + \left(\frac{C_2(\|h\|_\infty + \|h'\|_\infty)}{k}\right)\mathbb{E}\,|R'| \\ + \left(\frac{C_3(\|h\|_\infty + \|h'\|_\infty + \|h''\|_\infty)}{k}\right)\mathbb{E}\,|W'-W|^3. \end{aligned}$$

Construct an exchangable pair (W_n, W_n'), using Glauber dynamics, for the random variable

$$W_n = \frac{c_N}{n^{3/2}} \sum_{i,j=1}^{n} \langle\, \sigma_i, \sigma_j \rangle,$$

which is defined in section 3.2. We obtain

$$W_n' = W_n - \frac{c_N}{n^{3/2}} \sum_{j=1}^{n} \langle\, \sigma_I, \sigma_j \rangle + \frac{c_N}{n^{3/2}} \sum_{j=1}^{n} \langle \sigma_I', \sigma_j \rangle$$

The following lemma gives the bounds needed to apply Theorem 3.5.1 in this setting, and then Theorem 3.2.3 follows immediately.

Lemma 3.5.2. *For a fixed N, (W_n, W_n') as constructed above, $k = \frac{2c_N}{Nn^{3/2}}$ and $c = \frac{N}{(N+2)c_N^2}$, we have*

1. $\mathbb{E}\left[W_n' - W_n\big|\sigma\right] = Nk\left(1 - cW_n^2\right) + R$ *and* $\mathbb{E}\,|R| \leq \frac{C(log(n))}{n^2}$*;*
2. $\mathbb{E}\left[(W_n' - W_n)^2\big|\sigma\right] = kW_n + R'$*, and* $\mathbb{E}\,|R'| \leq \frac{C(log(n))}{n^2}$*;*
3. $\mathbb{E}\,|W_n' - W_n|^3 \leq \frac{C(log(n))}{n^{9/4}}$,

where C is a constant depending only on N, R and R' are defined below in the proof.

Proof of Lemma 3.5.2: For part (a), similar to [KM13],

$$\mathbb{E}\left[W_n' - W_n\big|\sigma\right] = -\frac{2}{n}W_n + \frac{2c_N}{n^{3/2}} + \frac{2c_N}{n^{5/2}} \sum_{i=1}^{n} f\left(\frac{N|\sigma^{(i)}|}{n}\right)|\sigma^{(i)}|,$$

where $f(\kappa) = \frac{I_{\frac{N}{2}}(\kappa)}{I_{\frac{N}{2}-1}(\kappa)}$. Near zero, $f(\kappa) \approx \frac{\kappa}{N} - \frac{\kappa^3}{N^2(N+2)}$, so substituting this value in the above equation yields

$$\sum_{i=1}^{n} f\left(\frac{N|\sigma^{(i)}|}{n}\right)|\sigma^{(i)}| = \sum_{i=1}^{n}\left[\frac{|\sigma^{(i)}|^2}{n} - \frac{N|\sigma^{(i)}|^4}{(N+2)n^3} + O\left(\frac{|\sigma^{(i)}|^6}{n^5}\right)\right]. \tag{3.12}$$

Note that

$$\frac{1}{n}\sum_{i=1}^{n}|\sigma^{(i)}|^2 = \frac{n^{3/2}W_n}{c_N} - \frac{2\sqrt{n}W_n}{c_N} + 1.$$

Similarly the second term on the R.H.S. of equation (3.12) is

$$-\frac{nNW_n^2}{(N+2)c_N^2} + \frac{4NW_n^2}{(N+2)c_N^2} - \frac{4N\sum_i\langle\, \sigma_i, S_n\rangle^2}{(N+2)n^3} - \frac{2NW_n}{(N+2)c_N\sqrt{n}} + \frac{NW_n}{(N+2)n^{3/2}c_N} + \frac{N}{(N+2)n^2}.$$

$$\mathbb{E}\left[W_n' - W_n\middle|\sigma\right] = \frac{Nc_N}{Nn^{3/2}} - \frac{N^2W_n^2}{c_NN(N+2)n^{3/2}} + R,$$

where

$$R = -\frac{2W_n}{n^2} + \frac{c_N}{n^{5/2}} + \frac{4NW_n^2}{(N+2)c_Nn^{5/2}} - \frac{4c_NN\sum_i\langle\, \sigma_i, S_n\rangle^2}{(N+2)n^{11/2}} - \frac{2NW_n}{(N+2)n^3} + \frac{NW_n}{(N+2)n^4} + \frac{c_NN}{(N+2)n^{9/2}},$$

Therefore we have,

$$\mathbb{E}\left[W_n' - W_n\middle|\sigma\right] = Nk\left(1 - cW_n^2\right) + R,$$

where $k = \frac{c_N}{Nn^{3/2}}$, $c = \frac{N}{(N+2)c_N^2}$ and $\mathbb{E}\,|R| \leq \frac{C(log(n))}{n^2}$.

This completes the proof.

For part (b), from the definition as before,

$$\mathbb{E}\left[(W_n' - W_n)^2\middle|\sigma\right] = \frac{c_N^2}{n^4}\sum_{i=1}^{n}\sum_{j,k\neq i}\mathbb{E}\left[\sigma_j^T(\sigma_i^* - \sigma_i)(\sigma_i^* - \sigma_i)^T\sigma_k\middle|\sigma, I = i\right]. \tag{3.13}$$

Using a previous computation, the terms $\mathbb{E}\left[(\sigma_i^* - \sigma_i)(\sigma_i^* - \sigma_i)^T\middle|\sigma, I = i\right]$ are

$$\left(\frac{a^2}{n^2}\right)\sum_{i=1}^{n}\left\{\left[\frac{I_{\frac{N}{2}}(\kappa) + \kappa I_{\frac{N}{2}+1}(\kappa)}{\kappa I_{\frac{N}{2}-1}(\kappa)}\right]P_i + \left[\frac{I_{\frac{N}{2}}(\kappa)}{\kappa I_{\frac{N}{2}-1}(\kappa)}\right]P_i^{\perp} - \left[\frac{I_{\frac{N}{2}}(\kappa)}{I_{\frac{N}{2}-1}(\kappa)}\right](r_i\sigma_i^T + \sigma_i r_i^T) + \sigma_i\sigma_i^T\right\}.$$

Again near zero, $f(\kappa) \approx \frac{\kappa}{N} - \frac{\kappa^3}{N^2(N+2)}$, so

$$\mathbb{E}\left[(\sigma_i^* - \sigma_i)(\sigma_i^* - \sigma_i)^T \middle| \sigma, I = i\right] = \frac{1}{N} Id - \frac{\kappa_i}{N}(r_i\sigma_i^T + \sigma_i r_i^T) + \sigma_i\sigma_i^T + R_i', \tag{3.14}$$

where

$$R_i' = \left(\frac{f(\kappa_i)}{\kappa_i} - \frac{1}{N}\right) Id - \left(f(\kappa_i) - \frac{\kappa_i}{N}\right)\left(r_i\sigma_i^T + \sigma_i r_i^T\right).$$

Ignoring the R_i' for the moment and putting the main term of (3.14) into (3.13) yields

$$\frac{c_N^2}{n^4}\sum_i \sum_{j,k\neq i}\left[\frac{1}{N}\langle\ \sigma_j, \sigma_k\rangle - \frac{\kappa_i}{N}(\sigma_j^T r_i \sigma_i^T \sigma_k + \sigma_j^T \sigma_i r_i^T \sigma_k) + \sigma_j^T \sigma_i \sigma_i^T \sigma_k\right].$$

$$\begin{aligned}
\mathbb{E}\left[(W_n' - W_n)^2 \middle| \sigma\right] &= \frac{c_N^2}{n^4}\sum_i \left[\frac{1}{N}|\sigma^{(i)}|^2 - \frac{2}{n}|\sigma^{(i)}|^2\langle\ \sigma_i, \sigma^{(i)}\rangle + \langle\ \sigma_i, \sigma^{(i)}\rangle^2\right] \\
&\quad + \frac{c_N^2}{n^4}\sum_i \sum_{j,k\neq i} \sigma_j^T R_i' \sigma_k \\
&= \frac{c_N^2}{n^4}\sum_i \left[\frac{2}{N}|\sigma^{(i)}|^2 - \frac{2}{n}|\sigma^{(i)}|^2\langle\ \sigma_i, \sigma^{(i)}\rangle\right] + \frac{c_N^2}{n^4}\sum_i \sum_{j,k\neq i} \sigma_j^T R_i' \sigma_k \\
&= \frac{2c_N^2}{Nn^3}\left(\frac{n^{3/2}W_n}{c_N} - \frac{2\sqrt{n}W_n}{c_N} + 1\right) - \frac{2c_N^2}{n^5}\sum_i |\sigma^{(i)}|^2\langle\ \sigma_i, \sigma^{(i)}\rangle
\end{aligned}$$

$$+\frac{c_N^2}{n^4}\sum_i \sum_{j,k\neq i} \sigma_j^T R_i' \sigma_k,$$

where the computation for $\mathbb{E}\left[\langle\ \sigma_i, \sigma^{(i)}\rangle^2\right]$ from the supercritical case has been used. Recall that the main term should be $\frac{1}{N}kW_n = \frac{c_N W_n}{Nn^{3/2}}$ and indeed it is. It is a routine collection of arguments very similar to those in the previous sections to show that the remaining terms are bounded in expectation by $\frac{C\log(n)}{n^2}$.

Finally, part (c) is similar to section 6 (c) of [KM13] with replaceing the unit sphere calculations with the corresponding unit hypersphere in N dimension.

$\square$

Chapter 4

Univariate exchangeable pair method for Subcritical Phase

In this chapter, we will give the proofs for lemma's of section 3.3 for the subcritical phase. The main idea is to use the series expansion for finding the conditional moments for the random variable $W_n' - W_n$. This chapter is an extended and more explanined version of the proofs for lemma 3.3.1 and lemma 3.3.2 presented in [KN16a].

4.1 Proof of Lemma 3.3.1

Let us denote $a := \sqrt{N-\beta}$.

For part (a), first define $f(\kappa) = \frac{I_{\frac{N}{2}}(\kappa)}{I_{\frac{N}{2}-1}(\kappa)}$. For $\beta < N$, we have $\frac{\beta|\sigma^{(i)}|}{n} = o(1)$ with probability exponentially close to 1. We therefore use the expansion of $f(c)$ near zero to write

$$\begin{aligned}\mathbb{E}\left[W_n' - W_n \middle| \sigma\right] = -\frac{1}{n}W_n + \left(\frac{a}{Nn^{5/2}}\sum_{i=1}^{n}\sum_{j\neq i}\beta\sigma_j\right) - \left(\frac{a}{N^2(N+2)n^{9/2}}\sum_{i=1}^{n}\sum_{j\neq i}\beta^3(\sigma_j)^3\right)\\ + \frac{a}{n^{3/2}}\sum_{i=1}^{n}\left[f(\kappa) - \frac{\kappa}{N} + \frac{\kappa^3}{N^2(N+2)}\right]\left(\frac{\sigma^{(i)}}{|\sigma^{(i)}|}\right).\end{aligned}$$

Since

$$\frac{a}{Nn^{5/2}}\sum_{i=1}^{n}\sum_{j\neq i}\beta\sigma_j = \frac{1}{Nn^2}\sum_{i=1}^{n}\left(\beta W_n - \frac{a\beta\sigma_i}{\sqrt{n}}\right) = \frac{1}{n}\left(\frac{\beta}{N} - \frac{\beta}{Nn}\right)W_n. \tag{4.1}$$

and

$$\frac{a}{N^2(N+2)n^{9/2}}\sum_{i=1}^{n}\sum_{j\neq i}\beta^3(\sigma_j)^3 = \frac{a\beta^3}{N^2(N+2)n^{9/2}}\left((n-1)+(n-2)(n-1)\,\mathbb{E}\langle\, \sigma_1,\sigma_2\rangle\right)\sum_{i=1}^{n}\sum_{j\neq i}\sigma_j.$$

Rewriting the right hand side of the last equation:

$$\left[\frac{a\beta}{Nn^{5/2}}\sum_{i=1}^{n}\sum_{j\neq i}\sigma_j\right]\left[\frac{\beta^2}{N(N+2)n^2}\left((n-1)+(n-2)(n-1)\,\mathbb{E}\langle\, \sigma_1,\sigma_2\rangle\right)\right].$$

Using (3.6) and (4.1) the last equation leads us to

$$\frac{a}{N^2(N+2)n^{9/2}}\sum_{i=1}^{n}\sum_{j\neq i}\beta^3(\sigma_j)^3 \approx \frac{\beta^3}{N^2(N+2)n^2}\left(W_n - \frac{W_n}{n}\right).$$

Therefore,

$$\begin{aligned}\mathbb{E}\left[W_n' - W_n\middle|\sigma\right] = -\frac{1}{n}W_n + \frac{W_n}{n}\left(\frac{\beta}{N} - \frac{\beta}{Nn}\right) - \frac{\beta^3}{N^2(N+2)n^2}\left(W_n - \frac{W_n}{n}\right)\\ + \frac{a}{n^{3/2}}\sum_{i=1}^{n}\left[f(\kappa) - \frac{\kappa}{N} + \frac{\kappa^3}{N^2(N+2)}\right]\left(\frac{\sigma^{(i)}}{|\sigma^{(i)}|}\right).\end{aligned}$$

$$\begin{aligned}\mathbb{E}\left[W_n' - W_n\middle|\sigma\right] = -\left(\frac{1}{n} - \frac{\beta}{Nn}\right)W_n - \frac{\beta}{Nn^2}W_n - \frac{\beta^3}{N^2(N+2)n^2}\left(W_n - \frac{W_n}{n}\right)\\ + \frac{a}{n^{3/2}}\sum_{i=1}^{n}\left[f(\kappa) - \frac{\kappa}{N} + \frac{\kappa^3}{N^2(N+2)}\right]\left(\frac{\sigma^{(i)}}{|\sigma^{(i)}|}\right).\end{aligned}$$

The matrix Λ of the theorem from [Mec09] is thus $\frac{N-\beta}{Nn}Id$ and

$$R = -\frac{\beta}{Nn^2}W_n - \frac{\beta^3}{N^2(N+2)n^2}\left(W_n - \frac{W_n}{n}\right) + \frac{a}{n^{3/2}}\sum_{i=1}^{n}\left[f(\kappa) - \frac{\kappa}{N} + \frac{\kappa^3}{N^2(N+2)}\right]\left(\frac{\sigma^{(i)}}{|\sigma^{(i)}|}\right).$$

This completes the proof of part(a).

For part (b), again similar to [KM13], we have:

$$\mathbb{E}\left[(W_n' - W_n)(W_n' - W_n)^T\middle|\sigma\right] = \frac{a^2}{n^2}\sum_{i=1}^{n}\frac{1}{Z_i}\int_{\mathbb{S}^{N-1}}(\theta-\sigma_i)(\theta-\sigma_i)^T\exp\left[\frac{\beta}{n}\sum_{j\neq i}\langle\ \sigma_j,\theta\rangle\right]d\mu(\theta).$$

Due to symmetry we have, $Z_i = Z_1$ for all i. Now letting $\theta = \theta_1 + \theta_2$, where θ_1 is the projection of θ onto the direction $\sigma^{(i)}$, the first term of the i^{th} summand is

$$\frac{1}{Z_1}\int_{\mathbb{S}^{N-1}}\left[\theta\theta^T\right]\exp\left[\frac{\beta}{n}\langle\ \sigma^{(i)},\theta\rangle\right]d\mu(\theta) = \frac{1}{Z_1}\int_{\mathbb{S}^{N-1}}\left[\theta_1\theta_1^T + \theta_2\theta_2^T\right]\exp\left[\frac{\beta}{n}\langle\ \sigma^{(i)},\theta\rangle\right]d\mu(\theta).$$

To compute it, write $r_i = \frac{\sigma^{(i)}}{|\sigma^{(i)}|}$, this implies $\theta_1 = \langle\ \theta, r_i\rangle r_i$, and $\theta_1\theta_1^T = |\langle\ \theta, r_i\rangle|^2\, r_i r_i^T$. Define $c := \frac{\beta|\sigma^{(i)}|}{n}$, Using the definition of Z_i we obtain

$$\begin{aligned}
\frac{1}{Z_1}\int_{\mathbb{S}^{N-1}} \theta_1\theta_1^T \exp\left[\frac{\beta}{n}\langle\ \sigma^{(i)},\theta\rangle\right] d\mu(\theta) &= \frac{1}{Z_1}\left(\int_{\mathbb{S}^{N-1}} |\langle\ \theta, r_i\rangle|^2 \exp\left[c\langle\ r_i,\theta\rangle\right] d\mu(\theta)\right) r_i r_i^T \\
&= \frac{1}{Z_1}\frac{A_{N-1}}{A_N}\left[\int_0^\pi \left(\cos(\theta_{N-1})\right)^2 e^{\kappa\cos(\theta_{N-1})}\sin^{N-2}(\theta_{N-1})d\theta_{N-1}\right] r_i r_i^T \\
&= \frac{I_{\frac{N}{2}}(\kappa)+\kappa I_{\frac{N}{2}+1}(\kappa)}{\kappa I_{\frac{N}{2}-1}(\kappa)} P_i.
\end{aligned}$$

$$\begin{aligned}
\frac{1}{Z_1}\int_{\mathbb{S}^{N-1}} \theta_1\theta_1^T \exp\left[\frac{\beta}{n}\langle\ \sigma^{(i)},\theta\rangle\right] d\mu(\theta) &= \frac{1}{Z_1}\left(\int_{\mathbb{S}^{N-1}} |\langle\theta, r_i\rangle|^2 \exp\left[c\langle\ r_i,\theta\rangle\right] d\mu(\theta)\right) r_i r_i^T \\
&= \frac{1}{Z_1}\frac{A_{N-1}}{A_N}\left[\int_0^\pi \left(\cos(\theta_{N-1})\right)^2 e^{\kappa\cos(\theta_{N-1})}\sin^{N-2}(\theta_{N-1})d\theta_{N-1}\right] r_i r_i^T \\
&= \frac{1}{Z_1}\frac{A_{N-1}}{A_N}\left[\int_{-1}^1 u^2 e^{\kappa u}\left(1-u^2\right)^{\frac{N-3}{2}} du\right] r_i r_i^T.
\end{aligned}$$

$$\frac{1}{Z_1}\int_{\mathbb{S}^{N-1}} \theta_1\theta_1^T \exp\left[\frac{\beta}{n}\langle\ \sigma^{(i)},\theta\rangle\right] d\mu(\theta) = \frac{I_{\frac{N}{2}}(\kappa)+\kappa I_{\frac{N}{2}+1}(\kappa)}{\kappa I_{\frac{N}{2}-1}(\kappa)} P_i,$$

where P_i is the orthogonal projection onto r_i. Let $\theta_2 = (x_1, x_2, ..., x_{N-1})$ be the orthonormal coordinate representation for θ_2 within $r_1^{\perp}$. Note that using symmetry, $i \neq j$ we have

$$\int_{\mathbb{S}^{N-1}} x_i x_j e^{c\langle\ r_i,\theta\rangle} d\mu(\theta) = 0$$

A polar coordinate expansion yields:

$$\begin{aligned}
&\int_{\mathbb{S}^{N-1}} x_{N-1}^2 e^{c\langle\ r_i,\theta\rangle} d\mu(\theta) \\
&= \frac{1}{A_N}\int_0^\pi ... \int_0^{2\pi} \left(\cos(\theta_{N-2})\sin(\theta_{N-1})\right)^2 e^{\kappa\cos(\theta_{N-1})} J_N d\theta_1 ... d\theta_{N-1} \\
&= \frac{A_{N-2}}{A_N}\int_0^\pi\int_0^\pi \left(\cos(\theta_{N-2})\sin(\theta_{N-1})\right)^2 e^{\kappa\cos(\theta_{N-1})}\sin^{N-3}(\theta_{N-2})\sin^{N-2}(\theta_{N-1})d\theta_{N-2}d\theta_{N-1} \\
&= \frac{A_{N-2}}{A_N}\frac{2\prod_{j=1}^{\frac{N}{2}-2}(2j)}{\prod_{j=0}^{\frac{N}{2}-1}(2j+1)}\int_0^\pi e^{\kappa\cos(\theta_{N-1})}\sin^N(\theta_{N-1})d\theta_{N-1} \\
&= \frac{A_{N-2}}{A_N}\frac{2\prod_{j=1}^{\frac{N}{2}-2}(2j)}{\prod_{j=0}^{\frac{N}{2}-1}(2j+1)}\left[\int_{-1}^1 e^{\kappa u}\left(1-u^2\right)^{\frac{N-1}{2}} du\right].
\end{aligned}$$

We notice by a series expansion that for $1 \leq j \leq N-1$, the values of $\int_{\mathbb{S}^{N-1}} x_j^2 e^{c\langle\ r_i,\theta\rangle} d\mu(\theta)$ are the same for very small κ. Therefore,

$$\frac{1}{Z_1}\int_{\mathbb{S}^{N-1}} \theta_2\theta_2^T \exp\left[\frac{\beta}{n}\langle\ \sigma^{(i)},\theta\rangle\right] d\mu(\theta) = \left[\frac{I_{\frac{N}{2}}(\kappa)}{\kappa I_{\frac{N}{2}-1}(\kappa)}\right] P_i^{\perp},$$

where $P_i^{\perp}$ is the orthogonal projection onto $r_1^{\perp}$.

$$\frac{1}{Z_1}\int_{\mathbb{S}^{N-1}} \theta\sigma_i^T \exp\left[\frac{\beta}{n}\langle\ \sigma^{(i)},\theta\rangle\right] d\mu(\theta) = \left[\frac{I_{\frac{N}{2}}(\kappa)}{I_{\frac{N}{2}-1}(\kappa)}\right] r_i\sigma_i^T.$$

$$\frac{1}{Z_1}\int_{\mathbb{S}^{N-1}} \theta^T\sigma_i \exp\left[\frac{\beta}{n}\langle\ \sigma^{(i)},\theta\rangle\right] d\mu(\theta) = \left[\frac{I_{\frac{N}{2}}(\kappa)}{I_{\frac{N}{2}-1}(\kappa)}\right] r_i\sigma_i^T.$$

$$\frac{1}{Z_1}\int_{\mathbb{S}^{N-1}} \sigma_i\sigma_i^T \exp\left[\frac{\beta}{n}\langle\ \sigma^{(i)},\theta\rangle\right] d\mu(\theta) = \sigma_i\sigma_i^T.$$

$$\begin{aligned}
&\mathbb{E}\left[(W_n'-W_n)(W_n'-W_n)^T\middle|\sigma\right]\\
&\quad= \frac{a^2}{n^2}\sum_{i=1}^n\left\{\left[\frac{I_{\frac{N}{2}}(\kappa)+\kappa I_{\frac{N}{2}+1}(\kappa)}{\kappa I_{\frac{N}{2}-1}(\kappa)}\right]P_i + \left[\frac{I_{\frac{N}{2}}(\kappa)}{\kappa I_{\frac{N}{2}-1}(\kappa)}\right]P_i^{\perp} - \left[\frac{I_{\frac{N}{2}}(\kappa)}{I_{\frac{N}{2}-1}(\kappa)}\right](r_i\sigma_i^T+\sigma_i r_i^T)+\sigma_i\sigma_i^T\right\}.
\end{aligned}$$

Remembering that $\kappa = \frac{\beta|\sigma^{(i)}|}{n}$, we have

$$\begin{aligned}
\mathbb{E}\left[(W_n'-W_n)(W_n'-W_n)^T\middle|\sigma\right] &= \frac{a^2}{n^2}\sum_{i=1}^n\left\{\frac{1}{N}P_i+\frac{1}{N}P_i^{\perp}-\frac{\kappa}{N}(r_i\sigma_i^T+\sigma_i r_i^T)+\sigma_i\sigma_i^T\right\}+R''\\
&= \frac{a^2}{Nn}Id+\frac{a^2}{Nn^2}\sum_{i=1}^n N\sigma_i\sigma_i^T - \frac{a^2\kappa}{Nn^2}\sum_{i=1}^n(r_i\sigma_i^T+\sigma_i r_i^T)+R'',
\end{aligned} \tag{4.2}$$

where the remainder term R'' is given by

$$\begin{aligned}
R'' = \frac{a^2}{n^2}\sum_{i=1}^n\Bigg\{&\left[\frac{I_{\frac{N}{2}}(\kappa)+\kappa I_{\frac{N}{2}+1}(\kappa)}{\kappa I_{\frac{N}{2}-1}(\kappa)}-\frac{1}{N}\right]P_i+\left[\frac{I_{\frac{N}{2}}(\kappa)}{cI_{\frac{N}{2}-1}(\kappa)}-\frac{1}{N}\right]P_i^{\perp}\\
&-\left[\frac{I_{\frac{N}{2}}(\kappa)}{I_{\frac{N}{2}-1}(\kappa)}-\frac{\kappa}{N}\right](r_i\sigma_i^T+\sigma_i r_i^T)\Bigg\}.
\end{aligned}$$

Using $r_i = \frac{\sigma^{(i)}}{|\sigma^{(i)}|}$ and $\kappa = \frac{\beta|\sigma^{(i)}|}{n}$, the third term of (4.2) simplifies:

$$\frac{a^2\kappa}{Nn^2}\sum_{i=1}^n(r_i\sigma_i^T+\sigma_i r_i^T) = \frac{a^2\beta}{Nn^3}\sum_{i=1}^n\sum_{j\neq i}(\sigma_j\sigma_i^T+\sigma_i\sigma_j^T) = \frac{2\beta}{Nn^2}W_nW_n^T - \frac{2a^2\beta}{Nn^3}\sum_{i=1}^n\sigma_i\sigma_i^T.$$

Collecting all terms,

$$
\begin{aligned}
&\mathbb{E}\left[(W_n' - W_n)(W_n' - W_n)^T \middle| \sigma\right] \\
&\quad = \left(\frac{N-\beta}{Nn}\right)\left[Id + \frac{1}{n}\sum_{i=1}^{n} N\sigma_i\sigma_i^T\right] - \frac{2\beta}{Nn^2}W_nW_n^T + \frac{2a^2\beta}{Nn^3}\sum_{i=1}^{n}\sigma_i\sigma_i^T + R'' + \frac{a^2}{Nn}Id - \frac{a^2}{Nn}Id \\
&\quad = 2\left(\frac{N-\beta}{Nn}\right)Id + R' \\
&\quad = 2\Lambda + R',
\end{aligned}
$$

where

$$
R' = \frac{a^2}{Nn}\left[\frac{1}{n}\sum_{i=1}^{n} N\sigma_i\sigma_i^T - Id\right] - \left[\frac{2\beta}{Nn^2}W_nW_n^T - \frac{2a^2\beta}{Nn^3}\sum_{i=1}^{n}\sigma_i\sigma_i^T\right] + R''
$$

□

4.2 Proof of Lemma 3.3.2

Recall that $\Lambda = \left(\frac{N-\beta}{Nn}\right) Id$, and thus $\|\Lambda^{-1}\|_{op} = \frac{Nn}{N-\beta}$. Now, from Lemma 3.3.1,

$$
R = -\frac{\beta}{Nn^2}W_n - \frac{\beta^3}{N^2(N+2)n^2}\left(W_n - \frac{W_n}{n}\right) + \frac{a}{n^{3/2}}\sum_{i=1}^{n}\left[f(\kappa) - \frac{\kappa}{N} + \frac{\kappa^3}{N^2(N+2)}\right]\left(\frac{\sigma^{(i)}}{|\sigma^{(i)}|}\right).
$$

It is important to note that we are considering first two terms of the series expansion of $f(\kappa)$ instead of only first term as was used in [KM13]. From our earlier heuristic approach, $\mathbb{E}\,|W_n|^2 \approx N$. We can use the same argument together with the fact that $\frac{I_{\frac{N}{2}}(\kappa)}{I_{\frac{N}{2}-1}(\kappa)} \leq \frac{\kappa}{N} - \frac{\kappa^3}{N^2(N+2)}$ to prove that $\mathbb{E}\,|W_n|^2 \leq N$. Using this condition we can bound the first two terms on R.H.S. of R as follows:

$$
\mathbb{E}\left[-\frac{\beta}{Nn^2}W_n - \frac{\beta^3}{N^2(N+2)n^2}\left(W_n - \frac{W_n}{n}\right)\right] \leq \left(-\frac{\beta}{N} - \frac{\beta^3}{N^2(N+2)}\left(1 - \frac{1}{n}\right)\right)\frac{\sqrt{N}}{n^2}
$$

For third term estimation of R, fix $\epsilon = \epsilon(n) \in (0,1)$ which will be defined later. Define $\widetilde{r}(\kappa) := \frac{I_{\frac{N}{2}}(\kappa)}{I_{\frac{N}{2}-1}(\kappa)} - \frac{\kappa}{N} + \frac{\kappa^3}{N^2(N+2)}$; observe that for b a universal constant, if $t \leq \epsilon$ then $|\widetilde{r}(t)| < b\epsilon^4$. Therefore

$$
\frac{a}{n^{3/2}}\left|\sum_{i=1}^{n}\left[\widetilde{r}\left(\frac{\beta|\sigma^{(i)}|}{n}\right)\right]\left(\frac{\sigma^{(i)}}{|\sigma^{(i)}|}\right)\right| \leq \frac{ba\epsilon^4}{\sqrt{n}} + \frac{a}{n^{3/2}}\sum_{i=1}^{n}\mathbb{1}\left(\frac{\beta|\sigma^{(i)}|}{n} > \epsilon\right), \tag{4.3}
$$

where we used $\left|\widetilde{r}\left(\frac{\beta|\sigma^{(i)}|}{n}\right)\right| \leq 1$ for any configuration σ. From an adaptation of proposition 3.1.5, since

$M_n = \frac{\sigma^{(i)}+\sigma_i}{n}$ the LDP for M_n, we deduce

$$\mathbb{P}\left[\frac{\beta|\sigma^{(i)}|}{n} > \epsilon\right] \leq C\exp\left[-\frac{n}{2}\inf\{I_\beta(r) : r \geq \epsilon\}\right],$$

where for $r = |x|$, from proposition 3.1.5 and (3.4) we have

$$I_\beta(r) = r\frac{I_{\frac{N}{2}}(r)}{I_{\frac{N}{2}-1}(r)} + \log\left[\frac{A_N}{A_{N-1}}\frac{r^{\frac{N}{2}-1}}{B_N\pi I_{\frac{N}{2}-1}(r)}\right] - \frac{\beta}{2}\left(\frac{I_{\frac{N}{2}}(r)}{I_{\frac{N}{2}-1}(r)}\right)^2,$$

where B_N is defined in (3.3). Using Taylor series expansion, we can deduce that there is a universal constant $q > 0$ such that for $\epsilon \in (0,1)$, $I_\beta(\epsilon) \geq \frac{\epsilon^2}{2N}\left(1 - \frac{\beta}{N}\right) - q\epsilon^3$. It follows that

$$\mathbb{P}\left[\frac{\beta|\sigma^{(i)}|}{n} > \epsilon\right] \leq C\exp\left[-\frac{n}{2}\left(\frac{\epsilon^2}{2N}\left(1 - \frac{\beta}{N}\right) - q\epsilon^3\right)\right].$$

Choose $\epsilon = \epsilon(n)$ such that $\epsilon^2 = \frac{4N\log(n)}{n\left(1-\frac{\beta}{N}\right)}$. Then $\mathbb{P}\left[\frac{\beta|\sigma^{(i)}|}{n} > \epsilon\right] \leq \frac{C'}{n}$, from the bound in (4.3) we notice that the second term is bounded by $n^{-3/2}$. Now we would be interested in bounding the first term. Notice that

$$\lim_{n\to\infty} n\epsilon^4 = \lim_{n\to\infty}\frac{16N^2\log(n)^2}{n\left(1 - \frac{\beta}{N}\right)^2} = 0$$

which implies that ϵ^4 is bounded above by $\frac{1}{n}$. This leads us to the conclusion that the first term of (4.3) is also bounded by $n^{-3/2}$. Therefore,

$$\frac{a}{n^{3/2}}\,\mathbb{E}\left|\sum_{i=1}^{n}\left[\widetilde{r}\left(\frac{\beta|\sigma^{(i)}|}{n}\right)\right]\left(\frac{\sigma^{(i)}}{|\sigma^{(i)}|}\right)\right| \;\leq\; \frac{ba\epsilon^4}{\sqrt{n}} + \frac{a}{\sqrt{n}}\mathbb{P}\left[\frac{\beta|\sigma^{(1)}|}{n} > \epsilon\right] \;\leq\; \frac{c_\beta(N)}{n^{3/2}}.$$

This completes the proof of part (a).

For part (b), for $x \in \mathbb{R}^n$, $\|xx^T\|_{HS} = |x|^2$, and thus $\mathbb{E}\,\|\sigma_i\sigma_i^T\|_{HS} = \mathbb{E}\,|\sigma_i|^2 = 1$, also

$$\mathbb{E}\,\|W_nW_n^T\|_{HS} = \mathbb{E}\,|W_n|^2 \leq N.$$

The value of R' from Lemma 3.3.1 is given by

$$R' = \frac{a^2}{Nn}\left[\frac{1}{n}\sum_{i=1}^{n} N\sigma_i\sigma_i^T - Id\right] - \left[\frac{2\beta}{Nn^2}W_nW_n^T - \frac{2a^2\beta}{Nn^3}\sum_{i=1}^{n}\sigma_i\sigma_i^T\right]$$
$$+ \frac{a^2}{n^2}\sum_{i=1}^{n}\left\{\left[\frac{I_{\frac{N}{2}}(\kappa)+\kappa I_{\frac{N}{2}+1}(\kappa)}{\kappa I_{\frac{N}{2}-1}(\kappa)} - \frac{1}{N}\right]P_i + \left[\frac{I_{\frac{N}{2}}(\kappa)}{cI_{\frac{N}{2}-1}(\kappa)} - \frac{1}{N}\right]P_i^{\perp}\right.$$
$$\left. - \left[\frac{I_{\frac{N}{2}}(\kappa)}{I_{\frac{N}{2}-1}(\kappa)} - \frac{\kappa}{N}\right](r_i\sigma_i^T + \sigma_i r_i^T)\right\}.$$

For estimating $\mathbb{E}\,\|\frac{1}{n}\sum_{i=1}^{n}(N\sigma_i\sigma_i^T - Id)\|_{HS}$, by the Cauchy-Schwarz inequality,

$$\mathbb{E}\left\|\frac{1}{n}\sum_{i=1}^{n}(N\sigma_i\sigma_i^T - Id)\right\|_{HS} \le \frac{1}{n}\sqrt{\sum_{i,j=1}^{n}\mathbb{E}\,Tr\big[(N\sigma_i\sigma_i^T - Id)(N\sigma_j\sigma_j^T - Id)\big]}.$$

So,

$$\mathbb{E}\,Tr\big[(N\sigma_i\sigma_i^T - Id)^2\big] = N^2\,\mathbb{E}\,|\sigma_i|^4 - 2N\,\mathbb{E}\,|\sigma_i|^2 + 2 = N^2 - 2N + N.$$

Similarly, for $i \neq j$,

$$\mathbb{E}\,Tr\big[(N\sigma_i\sigma_i^T - Id)(N\sigma_j\sigma_j^T - Id)\big] = N^2\,\mathbb{E}\left[\langle\, \sigma_i, \sigma_j\rangle^2\right] - N.$$

Using similar approach as in [KM13], we have:

$$\mathbb{E}\left[\langle\, \sigma_1, \sigma_2\rangle^2 \big| \{\sigma_i\}_{i\neq 1}\right] = \sigma_2^T\left(\left[\frac{I_{\frac{N}{2}}(\kappa)+\kappa I_{\frac{N}{2}+1}(\kappa)}{\kappa I_{\frac{N}{2}-1}(\kappa)}\right]P_i + \left[\frac{I_{\frac{N}{2}}(\kappa)}{\kappa I_{\frac{N}{2}-1}(\kappa)}\right]P_i^{\perp}\right)\sigma_2,$$

Again since $\kappa = o(1)$ with high probability, $\frac{I_{\frac{N}{2}}(\kappa)}{I_{\frac{N}{2}-1}(\kappa)} \le \frac{\kappa}{N} - \frac{\kappa^3}{N^2(N+2)}$ and

$$\mathbb{E}\left[\langle\, \sigma_1, \sigma_2\rangle^2 \big| \{\sigma_i\}_{i\neq 1}\right] \approx \sigma_2^T\left(\left(\frac{1}{N} + \frac{(N-1)\kappa^2}{N^2(N+2)}\right)P_i + \left(\frac{1}{N} - \frac{\kappa^2}{N^2(N+2)}\right)P_i^{\perp}\right)\sigma_2,$$

$$\mathbb{E}\left[\langle\, \sigma_1, \sigma_2\rangle^2 \big| \{\sigma_i\}_{i\neq 1}\right] \approx \frac{1}{N} - \frac{\kappa^2}{N^2(N+2)} + \frac{N\kappa^2}{N^2(N+2)}\,\mathbb{E}\left[\sigma_2^T\sigma_1\sigma_1^T\sigma_2 \big| \{\sigma_i\}_{i\neq 1}\right],$$

$$\mathbb{E}\left[\langle\ \sigma_1,\sigma_2\rangle^2\right]\approx\frac{1}{N}+\mathbb{E}\left[-\frac{\kappa^2}{N^2(N+2)}+\frac{N\kappa^2}{N^2(N+2)}\langle\ \sigma_1,\sigma_2\rangle^2\right].$$

This implies

$$\mathbb{E}\,Tr\big[(N\sigma_i\sigma_i^T-Id)(N\sigma_j\sigma_j^T-Id)\big]\approx-\frac{\kappa^2}{(N+2)}+\frac{N\kappa^2}{(N+2)}\,\mathbb{E}\left[\langle\ \sigma_1,\sigma_2\rangle^2\right].$$

For each pair $i\neq j$ and r_1 being a universal constant, the error in this approximation is represented by

$$\mathbb{E}\left|N^2\sigma_2^T\left(\left[\frac{I_{\frac{N}{2}}(\kappa)+\kappa I_{\frac{N}{2}+1}(\kappa)}{\kappa I_{\frac{N}{2}-1}(\kappa)}-\frac{1}{N}-\frac{(N-1)\kappa^2}{N^2(N+2)}\right]P_i+\left[\frac{I_{\frac{N}{2}}(\kappa)}{\kappa I_{\frac{N}{2}-1}(\kappa)}-\frac{1}{N}+\frac{\kappa^2}{N^2(N+2)}\right]P_i^{\perp}\right)\sigma_2\right|$$

and it is bounded by $r_1\,\mathbb{E}\,\kappa^3$. Now,

$$\mathbb{E}\,\kappa^2=\frac{\beta^2}{n^2}\sum_{i,j>1}\mathbb{E}\langle\ \sigma_i,\sigma_j\rangle\leq\frac{\beta^2}{n^2}\left[(n-1)+(n-1)(n-2)\frac{\beta}{n(N-\beta)}\right]\leq\frac{N\beta^2}{n(N-\beta)},$$

and

$$\mathbb{E}\,c^3\leq\frac{N\beta^3}{n(N-\beta)},$$

therefore

$$\mathbb{E}\left\|\frac{1}{n}\sum_{i=1}^{n}(N\sigma_i\sigma_i^T-Id)\right\|_{HS}\leq\sqrt{\frac{(N^2-2N+N)+\frac{r_2(\beta^2+\beta^3)}{(N-\beta)}}{n}}.$$

Using Taylor expansion for R'' (which is remaining term of the error R' in Lemma 3.3.1) and for a universal constant c_*,we obtain:

$$\begin{aligned}\mathbb{E}\,\|R''\|_{HS}\leq\frac{a^2}{n^2}\sum_{i=1}^{n}\mathbb{E}\Bigg\{&\left\|\left[\frac{I_{\frac{N}{2}}(\kappa)+\kappa I_{\frac{N}{2}+1}(\kappa)}{\kappa I_{\frac{N}{2}-1}(\kappa)}-\frac{1}{N}\right]P_i\right\|_{HS}+\left\|\left[\frac{I_{\frac{N}{2}}(\kappa)}{\kappa I_{\frac{N}{2}-1}(\kappa)}-\frac{1}{N}\right]P_i^{\perp}\right\|_{HS}\\&-\left\|\left[\frac{I_{\frac{N}{2}}(\kappa)}{I_{\frac{N}{2}-1}(\kappa)}-\frac{\kappa}{N}\right](r_i\sigma_i^T+\sigma_ir_i^T)\right\|_{HS}\Bigg\}\\\leq\frac{c_*\sqrt{N(N-\beta)}\beta}{n^{3/2}}\leq\frac{c_\beta(N)}{n^{3/2}},&\end{aligned}$$

where we used the facts that $\|P_i\|_{HS}$, $\|P_i^{\perp}\|_{HS}$ and $\|r_i\sigma_i^T\|_{HS}$ are all bounded by $\sqrt{N}$ or smaller and that $\mathbb{E}\,|\sigma^{(i)}|\leq\sqrt{\frac{Nn}{N-\beta}}$.This completes the proof of part (b).

Finally, part (c) is trivial and identical to [KM13] with different σ belonging to the sphere.

$\square$

Chapter 5

Application of Stein's method

In this chapter, we will derive the expressions for the free energy, Stein's characteristic operator and will also give the proof for the density function for the mean-field $O(N)$ models. The results discussed here are recently published in [KN16a].

5.1 Free energy

We will start this section by revisiting that the free energy can be obtained by minimizing the functional

$$\Phi_\beta(r) = r\frac{I_{\frac{N}{2}}(r)}{I_{\frac{N}{2}-1}(r)} + \log\left[\frac{A_N}{A_{N-1}}\frac{r^{\frac{N}{2}-1}}{B_N \pi I_{\frac{N}{2}-1}(r)}\right] - \frac{\beta}{2}\left(\frac{I_{\frac{N}{2}}(r)}{I_{\frac{N}{2}-1}(r)}\right)^2.$$

where B_N is defined in (3.3) and $r = g^{-1}(\beta)$, with

$$g(r) = g_N(r) := r\frac{I_{\frac{N}{2}-1}(r)}{I_{\frac{N}{2}}(r)}$$

Lemma 5.1.1. *Consider the functional defined above:*

1. *For $\beta \le N$, the $\inf_{b\ge 0}\{\Phi_\beta(b)\} = 0$ achieved only at $b = 0$.*
2. *For $\beta > N$, there is a unique value of $r \in (0,\infty)$ which minimizes Φ_β over $[0,\infty)$.*
3. *Let b denote the unique positive solution of $b - \beta f(b) = 0$ with*

$$f(b) = \frac{I_{\frac{N}{2}}(b)}{I_{\frac{N}{2}-1}(b)}$$

Then $\beta f'(b) < 1$. In particular, $\Phi'_\beta(b) = 0$ and $\Phi''_\beta(b) > 0$.

Proof of Lemma 5.1.1:

1. We first prove that $\Phi_\beta(x)$ is increasing on $(0,\infty)$ for $\beta \le N$. Taking derivative of $\Phi_\beta(x)$ and using the recursive relation for modified Bessel function of first kind we have:

$$\Phi'_\beta(b) = \left(b - \beta\frac{I_{\frac{N}{2}}(b)}{I_{\frac{N}{2}-1}(b)}\right)\left(1 - \frac{N-1}{b}\frac{I_{\frac{N}{2}}(b)}{I_{\frac{N}{2}-1}(b)} - \left(\frac{I_{\frac{N}{2}}(b)}{I_{\frac{N}{2}-1}(b)}\right)^2\right).$$

The positivity of second expression in last equation is already proved in [Amo74a]. Therefore, in order to prove that $\Phi_\beta(b)$ is increasing on $(0,\infty)$ our problem reduced to proving that

$$\frac{I_{\frac{N}{2}}(b)}{I_{\frac{N}{2}-1}(b)} < \frac{b}{\beta}$$

Since $\beta \leq N$, therefore we just need to show that above expression holds for $\beta = N$. Let $\nu + 1 = \frac{N}{2}$, then the last expression simplifies to

$$\frac{I_{\nu+1}(b)}{I_\nu(b)} < \frac{b}{2(\nu+1)}.$$

The last inequality is already proved by Ifantis and Siafarikas [IS90]. Also,

$$\lim_{b\to 0}\frac{I_{\frac{N}{2}}(b)}{I_{\frac{N}{2}-1}(b)} = 0,$$

$$\lim_{b\to 0}\log\left[\frac{A_N}{A_{N-1}}\frac{b^{\frac{N}{2}-1}}{B_N\pi I_{\frac{N}{2}-1}(b)}\right] = 0.$$

Therefore,

$$\lim_{b\to 0}\Phi_\beta(b) = 0.$$

2. Expanding $\Phi'_\beta(b)$ near $x = 0$, we have

$$\Phi'_\beta(b) = \left(b - \beta\frac{I_{\frac{N}{2}}(b)}{I_{\frac{N}{2}-1}(b)}\right)\left(1 - \frac{N-1}{b}\frac{I_{\frac{N}{2}}(b)}{I_{\frac{N}{2}-1}(b)} - \left(\frac{I_{\frac{N}{2}}(b)}{I_{\frac{N}{2}-1}(b)}\right)^2\right) \approx \frac{(N-\beta)x}{N^2}.$$

This implies that for $\beta > N$, $x = 0$ is a local maximum of $\Phi_\beta(b)$ on $[0,\infty)$. Since

$$\left(1 - \frac{N-1}{b}\frac{I_{\frac{N}{2}}(b)}{I_{\frac{N}{2}-1}(b)} - \left(\frac{I_{\frac{N}{2}}(b)}{I_{\frac{N}{2}-1}(b)}\right)^2\right) > 0,$$

using similar reasoning as [KM13], at the interior minimum we have

$$\beta = b\frac{I_{\frac{N}{2}-1}(b)}{I_{\frac{N}{2}}(b)}.$$

Simpson and Spector[SS84] have already proved that, for all N, above expression is strictly increasing and convex. Therefore, the above equation uniquely defines b in terms of β.

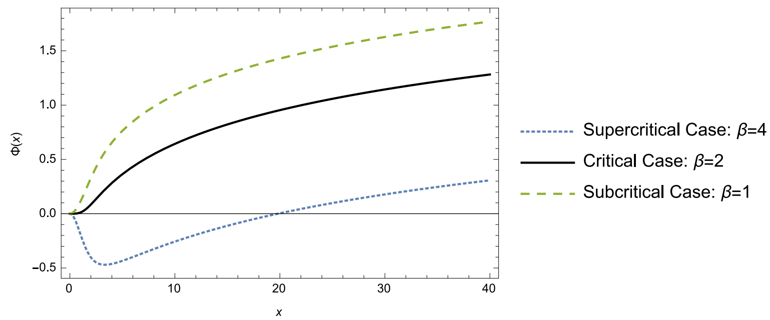

Figure 5.1: Graphical representation of functional $\Phi_\beta(x)$ for the mean-field XY Model

3. Let

$$f(x) = x - \beta \frac{I_{\frac{N}{2}}(x)}{I_{\frac{N}{2}-1}(x)},$$

such that $f(b) = 0$. Again using similar reasoning as in [KM13] , there is an $x < b$ such that $f'(x) = 0$. Now

$$f'(x) = 1 - \beta \left(1 - \frac{N-1}{x} \frac{I_{\frac{N}{2}}(x)}{I_{\frac{N}{2}-1}(x)} - \left(\frac{I_{\frac{N}{2}}(x)}{I_{\frac{N}{2}-1}(x)} \right)^2 \right)$$

Consider $g(x) := \frac{N-1}{x} \frac{I_{\frac{N}{2}}(x)}{I_{\frac{N}{2}-1}(x)} + \left(\frac{I_{\frac{N}{2}}(x)}{I_{\frac{N}{2}-1}(x)} \right)^2$, then

$$\begin{aligned} g'(x) &= (1 - g(x)) \left(\frac{2\nu - 1}{x} + 2\frac{I_{\frac{N}{2}}(x)}{I_{\frac{N}{2}-1}(x)} \right) - \frac{2\nu - 1}{x^2} \frac{I_{\frac{N}{2}}(x)}{I_{\frac{N}{2}-1}(x)} \\ &> (1 - g(x)) \left(\frac{2\nu - 1}{x} + 2\frac{I_{\frac{N}{2}}(x)}{I_{\frac{N}{2}-1}(x)} \right) - \frac{2\nu - 1}{x} (1 - g(x)) \\ &> 2 (1 - g(x)) \frac{I_{\frac{N}{2}}(x)}{I_{\frac{N}{2}-1}(x)} \\ &> \ 0, \end{aligned}$$

where in the second step we have used the following improved bound proved in page 243 of [Amo74a],

$$1 - \frac{N-1}{x} \frac{I_{\frac{N}{2}}(x)}{I_{\frac{N}{2}-1}(x)} - \left(\frac{I_{\frac{N}{2}}(x)}{I_{\frac{N}{2}-1}(x)} \right)^2 < \frac{1}{x} \frac{I_{\frac{N}{2}}(x)}{I_{\frac{N}{2}-1}(x)}.$$

Therefore, $1 - g(x)$ is decreasing on $(0, \infty)$ and $f'(x)$ is increasing on $(0, \infty)$ and there is only one zero of f' with $f'(b) > 0$. Now using part (2) we have,

$$\Phi''_\beta(b) = (1 - \beta (1 - g(b))) (1 - g(b)) = f'(b) (1 - g(b)) > 0$$

$\square$

5.2 Stein's Characteristic Operator

It is to be noted that for applications of Stein's method we need to identify the characteristic operator and density of the distributions.

Lemma 5.2.1. *A random variable $Y > 0$ has density*

$$p(t) = \begin{cases} \frac{1}{Z} t^{\frac{N-2}{2}} e^{-\widetilde{k}t^2} & t \geq 0 \\ 0 & t < 0 \end{cases}$$

if and only if

$$\mathbb{E}\left[Y f'(Y) + \left(\frac{N}{2} - 2\widetilde{k}Y^2\right) f(Y)\right] = 0 \tag{5.1}$$

for $\widetilde{k} = \frac{1}{N^2(4N+8)}$ and all $f \in C^1((0,\infty))$ such that $\int_0^\infty f(t)p(t)dt < \infty$. The corresponding related distribution has a characterizing operator T_p which is invertible on space $\{h : \mathbb{E}\, h(x) = 0\}$ and defined by:

$$[T_p f](x) = xf'(x) + \left(\frac{N}{2} - 2\widetilde{k}x^2\right) f(x).$$

Proof of Lemma 5.2.1:

Consider a positive random variable Y which has $p(t)$ as its density function, then using integration by parts it is straightforward to show that Y satisfies (5.1).

Conversely, consider a random variable X having $p(t)$ as its density function. Then given $h : (0,\infty) \to \mathbb{R}$, we construct $f = f_h$ so that

$$tf'(t) + \left(\frac{N}{2} - 2\widetilde{k}t^2\right) f(t) = h(t) - \mathbb{E}\, h(X).$$

We claim that the solution f is given by

$$\begin{aligned} f(t) &= \frac{1}{tp(t)} \int_0^t \big[h(s) - \mathbb{E}\, h(X)\big] p(s) ds \\ &= -\frac{1}{tp(t)} \int_t^\infty \big[h(s) - \mathbb{E}\, h(X)\big] p(s) ds. \end{aligned}$$

To see this, we differentiate the expression similar to [KM13] and deduce that

$$h(t) - \mathbb{E}\, h(X) = f(t) + tf'(t) + \frac{tf(t)p'(t)}{p(t)} = \left(\frac{N}{2} - 2\widetilde{k}t^2\right) f(t) + tf'(t).$$

Therefore for bounded f and f' and Y satisfying (5.1), then for given h, $f = f_h$ solves the Stein equation,

$$\mathbb{E}\, h(Y) - \mathbb{E}\, h(X) = \mathbb{E}\left[Yf'(Y) + \left(\frac{N}{2} - 2\widetilde{k}Y^2\right) f(Y)\right] = 0,$$

thus $Y \stackrel{d}{=} X$. □

Lemma 5.2.2. *The characteristic operator defined above has the following boundedness results: Let $h : \mathbb{R} \to \mathbb{R}$ be given. Suppose that*

$$f(t) = f_h(t) := \frac{1}{tp(t)} \int_0^t \big[h(s) - \mathbb{E}\, h(X)\big] p(s) ds,$$

with p as defined in the previous lemma and $\widetilde{k} = \frac{1}{N^2(4N+8)}$. Then $[T_p f_h](x) = h(x) - \mathbb{E}\, h(X)$ and

1. *$\|f_h\|_\infty \leq G(\frac{N}{4})\|h\|_\infty$, where $G(x) = \left(\frac{e}{x}\right)^x (\Gamma(x) - \Gamma(x, x))$.*

2. *$\|f_h'\|_\infty \leq \big(2\ G(\frac{N}{4}) + N\big) \widetilde{k} t_0 \|h\|_\infty + \frac{N}{4}\|h'\|_\infty$.*

3. *$\|f_h''\|_\infty \leq K_1\|h\|_\infty + K_2\|h'\|_\infty + K_3\|h''\|_\infty$, where K_1, K_2, K_3 are constants depending on dimension N and c_N.*

Proof of Lemma 5.2.2:

1. The first formula for f_h, gives the following bound:

$$f(t) \leq \frac{2\|h\|_\infty}{tp(t)} \left(\int_0^t p(s) ds\right).$$

For $t \leq t_0 = \sqrt{\frac{N}{4\widetilde{k}}}$, we have

$$\begin{aligned}
\frac{\int_0^t p(s)ds}{tp(t)} &= \frac{1}{2}\left(\frac{2N}{t}\right)^{\frac{N}{2}} e^{\widetilde{k}t^2}(N+2)^{\frac{N}{4}}\left(\Gamma\left(\frac{N}{4}\right) - \Gamma\left(\frac{N}{4}, t^2\,\widetilde{k}\right)\right) \\
&\leq \frac{1}{2}\left(\frac{4e}{N}\right)^{\frac{N}{4}}\left(\Gamma\left(\frac{N}{4}\right) - \Gamma\left(\frac{N}{4}, \frac{N}{4}\right)\right),
\end{aligned}$$

Therefore we have,

$$f(t) \leq \left(\frac{4e}{N}\right)^{\frac{N}{4}}\left(\Gamma\left(\frac{N}{4}\right) - \Gamma\left(\frac{N}{4}, \frac{N}{4}\right)\right)\|h\|_\infty.$$

Also using the fact that X has density function $p(t)$ and from the definition of f, similar approach as in [KM13] can be used to find bound on $|f(t)|$. For any fixed N and $t \geq t_0$, we have

$$|f(t)| \leq \|h\|_\infty.$$

2. We know that f solves the Stein equation,

$$tf'(t) = \left(2\widetilde{k}t^2 - \frac{N}{2}\right)f(t) + h(t) - \mathbb{E}\,h(X).$$

For $t \leq t_0$, we have

$$\begin{aligned}
h(t) - \mathbb{E}\,h(X) - \frac{N}{2}f(t) &= h(t) - \mathbb{E}\,h(X) - \frac{N}{2tp(t)}\int_0^t [h(s) - \mathbb{E}\,h(X)]p(s)ds \\
&= \frac{N}{2tp(t)}\int_0^t \Big([h(t) - \mathbb{E}\,h(X)]\left(\frac{s^{\frac{N-2}{2}}}{t^{\frac{N-2}{2}}}\right)p(t) - [h(s) - \mathbb{E}\,h(X)]p(s)\Big)ds.
\end{aligned}$$

Now, notice that $\frac{p(s)}{p(t)} \leq 1$ so we have

$$\begin{aligned}
\left|\frac{1}{tp(t)}\int_0^t [h(t) - \mathbb{E}\,h(X)]\left(\left(\frac{s}{t}\right)^{\frac{N-2}{2}}p(t) - p(s)\right)ds\right| &\leq \frac{2\|h\|_\infty}{tp(t)}\int_0^t \left|1 - \left(\frac{s}{t}\right)^{\frac{N-2}{2}}\frac{p(t)}{p(s)}\right|p(s)ds \\
&= \frac{2\|h\|_\infty}{tp(t)}\int_0^t \left|1 - e^{\widetilde{k}(s^2 - t^2)}\right|p(s)ds \\
&\leq 2\widetilde{k}\|h\|_\infty t^2.
\end{aligned}$$

Also,

$$\left|\frac{1}{tp(t)}\int_0^t \Big([h(t) - \mathbb{E}\,h(X)] - [h(s) - \mathbb{E}\,h(X)]\Big)p(s)ds\right| \leq \frac{\|h'\|_\infty}{tp(t)}\int_0^t (t-s)p(s)ds \leq \frac{\|h'\|_\infty t}{2}.$$

This implies for $t \leq t_0$, we have

$$\frac{1}{t}\left|h(t) - \mathbb{E}\, h(X) - \frac{N}{2} f(t)\right| \leq N\widetilde{k} t_0 \|h\|_\infty + \frac{N}{4}\|h'\|_\infty,$$

and since

$$f'(t) = 2\widetilde{k} t f(t) + \frac{1}{t}\left(h(t) - \mathbb{E}\, h(X) - \frac{N}{2} f(t)\right),$$

as a result, we have

$$\begin{aligned}\left|f'(t)\right| &\leq 2\widetilde{k} t_0 \|f\|_\infty + N\widetilde{k} t_0 \|h\|_\infty + \frac{N}{4}\|h'\|_\infty \\ &\leq \left(2\; G\left(\frac{N}{4}\right) + N\right)\widetilde{k} t_0 \|h\|_\infty + \frac{N}{4}\|h'\|_\infty.\end{aligned}$$

For $t \geq t_0 = \sqrt{\frac{N}{4\widetilde{k}}}$, from Stein equation we get:

$$\begin{aligned}|f'(t)| &\leq 2\widetilde{k} t |f(t)| + \frac{\frac{N}{2}\|f\|_\infty + 2\|h\|_\infty}{t} \\ &\leq \frac{4\widetilde{k}\|h\|_\infty P\left[X \geq t\right]}{p(t)} + \frac{\frac{N}{2}\|f\|_\infty + 2\|h\|_\infty}{t}.\end{aligned}$$

Using the estimate

$$P\left[X \geq t\right] \leq \frac{\Gamma\left(\frac{N}{4}, \frac{t^2}{4N^2(N+2)}\right)}{\Gamma\left(\frac{N}{4}\right)}$$

along with some simplifications completes the proof.

3. Consider again the Stein equation

$$f'(t) = \left(2\widetilde{k} t - \frac{N}{2t}\right) f(t) + \frac{1}{t}\left(h(t) - \mathbb{E}\, h(X)\right).$$

differentiating both sides with respect to t and substituting value of $f'(t)$ from above, we obtain

$$\begin{aligned}f''(t) &= 2\widetilde{k}\left(f(t) + t f'(t)\right) + \frac{N}{2t^2} f(t) - \frac{N}{2t} f'(t) - \frac{1}{t^2}\left(h(t) - \mathbb{E}\, h(X)\right) + \frac{h'(t)}{t} \\ &= 2\widetilde{k}\left(f(t) + t f'(t)\right) - \frac{N}{2t} f'(t) + \frac{1}{t}\left(h'(t) + \frac{\frac{N}{2} f(t) - \left[h(t) - \mathbb{E}\, h(X)\right]}{t}\right).\end{aligned}$$

Using a similar approach to [KM13], the last term from above simplifies to

$$\begin{aligned}
&h'(t) - \frac{[h(t) - \mathbb{E}\,h(X)] - \frac{N}{2}f(t)}{t} \\
&= h'(t) - \frac{N}{2t^2 p(t)} \int_0^t \left([h(t) - \mathbb{E}\,h(X)] \left(\frac{s}{t}\right)^{\frac{N-2}{2}} p(t) - [h(s) - \mathbb{E}\,h(X)] p(s) \right) ds \\
&= -\frac{N}{2t^2 p(t)} \int_0^t \left([h(t) - \mathbb{E}\,h(X)] \left(\frac{s}{t}\right)^{\frac{N-2}{2}} p(t) - [h(s) - \mathbb{E}\,h(X)] p(s) - \left(\frac{s}{t}\right)^{\frac{N-2}{2}} (t-s) h'(t) p(t) \right) ds \\
&= -\frac{N}{2t^2 p(t)} \int_0^t \left([h(t) - \mathbb{E}\,h(X)] - [h(s) - \mathbb{E}\,h(X)] - (t-s) h'(t) \right) \frac{e^{-\widetilde{k}t^2} s^{\frac{N-2}{2}}}{z} ds.
\end{aligned}$$

Define

$$H(t) = [h(t) - \mathbb{E}\,h(X)]\, p(t),$$

then

$$(t-s)h'(t) = (t-s)H'(t) + 2\widetilde{k}t[h(t) - \mathbb{E}\,h(X)]$$

Then the above simplifies to

$$\begin{aligned}
&h'(t) - \frac{\left[h(t) - \mathbb{E}\,h(X) - \frac{N}{2}f(t)\right]}{t} \\
&= \frac{2}{t^2 p(t)} \int_0^t \left(H(s) - H(t) - (s-t)H'(t) + 2\widetilde{k}t\,[h(t) - \mathbb{E}\,h(X)] \right) \frac{e^{-\widetilde{k}t^2} s^{\frac{N-2}{2}}}{z} ds.
\end{aligned}$$

The rest of this part is computing the bounds for all terms similar to [KM13], which we leave for the reader to check. $\square$

5.3 Proof of Theorem 3.5.1

In this section we will give the proof for Theorem 3.5.1 using above two lemmas.

Given h, let f be the solution to the Stein equation described above. Then by exchangeability and the conditions on (W, W'),

$$\begin{aligned}
0 &= \mathbb{E}[(W' - W)(f(W') + f(W))] \\
&= \mathbb{E}[(W' - W)(f(W') - f(W)) + 2(W' - W)f(W)] \\
&= \mathbb{E}[(W' - W)^2 f'(W) + E'' + 2Nk(1 - cW^2)f(W) + 2Rf(W)] \\
&= \mathbb{E}[kWf'(W) + R'f'(W) + E'' + 2Nk(1 - cW^2)f(W) + 2Rf(W)].
\end{aligned}$$

where

$$E'' = \sum_{n=2}^{\infty} \frac{f^n(W)}{n!}(W' - W)^{n+1}$$

Then

$$\mathbb{E}[Wf'(W) + 2N(1 - cW^2)f(W)] = -\frac{1}{k}\,\mathbb{E}[R'f'(W) + 2Rf(W) + E''],$$

and

$$|E''| \leq \frac{\|f''\|_\infty}{2}|(W' - W)|^3.$$

The result is thus immediate from Lemma 5.2.2. $\square$

5.4 Density Function Calculation

Consider the functional $G_\beta(r)$ defined as:

$$G_\beta(r) = r\frac{I_{\frac{N}{2}}(r)}{I_{\frac{N}{2}-1}(r)} + \log\left[\frac{A_N}{A_{N-1}}\frac{r^{\frac{N}{2}-1}}{B_N \pi I_{\frac{N}{2}-1}(r)}\right] - \frac{\beta}{2}\left(\frac{I_{\frac{N}{2}}(r)}{I_{\frac{N}{2}-1}(r)}\right)^2,$$

where B_N is defined in (3.3). We can check that $r = 0$ is point of inflection for $G(r)$ at the critical value $\beta = N$. We can write the Taylor expansion for the critical case as follows:

$$G(r) = G(m) + \lambda(m)\frac{(r-m)^{2k+1}}{(2k+1)!} + O(r)^5,$$

where $m = 0$, $k = \frac{3}{2}$, $G(0) = 0$ and $\lambda(0) = \frac{3!}{N^2(2+N)}$. Finally from Theorem 5 of [EN78b], our density function for $r \geq 0$ is given by

$$p(r) = \frac{1}{\tilde{z}}r^{N-1}e^{-\tilde{k}r^4},$$

with $\tilde{k} = \frac{1}{4N^2(N+2)}$. Using substitution $t = r^2$ we obtain:

$$p(t) = \frac{1}{z}t^{\frac{N-2}{2}}e^{-\tilde{k}t^2}.$$

Therefore, the density function at the critical temperature for the $O(N)$-model is given by

$$p(t) = \begin{cases} \frac{1}{z}t^{\frac{N-2}{2}}e^{-\tilde{k}t^2} & t \geq 0; \\ 0 & t < 0, \end{cases}$$

The reader can also verify the above density function using an approach similar to [SDHR04].

Chapter 6

Concluding Remarks

In chapters 1-5, we have studied the limit theorems along with the conditional moments corresponding to each phase. In statistical mechanics, we are also interested in calculating the physical properties of the complex systems using Monte Carlo methods (MCM). For spin-glass systems, at low temperature we observe that these systems settles into ground states (i.e, minimal energy states). The computational task here is to calculate the statistical properties of the system such as magnetization, susceptibility, critical temperature and average energy. The size of the spin-glass system problems is really large, therefore we use Monte Carlo methods to generate a sequence of spin configurations (hoping to get a sufficient representation of the original system) for statistical measurements.

In order to calculate these statistics,we need to find all the ground states effectively using Monte Carlo methods. At low temperature, the statistically important configurations are composed of big blocks of mostly aligned spins. Using single spin flips Monte Carlo Metropolis method it's really hard to change such big size blocks (i.e., these aligned spins blocks creates barrier which are hard to cross). This results in slow transition between such blocks. Also the fluctuations of the blocks of the spins at low temperature can not be eliminated using these single spin flips Monte Carlo Metropolis methods. The need for block-wise flips at low temperature suggests that multigrid Monte Carlo methods (MGMC) are a good candidate for solving such problems in these spin glass systems. MGMC methods were first proposed to deal with these issues [KDB89, KDR+88]. Later they were extended to calculate the thermodynamics quantities, such as average energy, magnetization etc accurately within a reasonable computer time [BGR94]. A detailed developmental explanation of MGMC methods for the Ising model, as well as XY model, is available in [Ron90], where a class of problems were solved using these methods by using different cluster algorithms. The main idea is to create the stochastic clusters whose flipping was decided based on whether the updated cluster will decrease the energy of the system or not (as we are working with the energy minimization problem). One of our current project is to investigate, for the $O(N)$ models in absense of external field, how Multigrid Monte Carlo Methods (MGMC) are used to solve the energy minimization problem at low temperature.

Another open question is to find the critical dimension for these $O(N)$ models.Critical dimension corresponds to the dimensionality of space at which the characteristics of the phase transition changes. Above the upper critical dimension the critical exponents of the lattice theory become the same as that in mean field theory. Based on my research project on asymptotics of $O(N)$ models [KN16a], now i would like to see how MGMC methods are useful in finding the critical dimension in these $O(N)$-models.

I am also interested in calculating the critical exponents of the mean-field $O(N)$ models. Critical exponents plays an important role in studying the behavior of the physical quantities near phase transition. One of the open problem is to prove the universality of the critical exponents i.e., they are independent of spin dimension, range of interaction. I believe i have some analytic results for this claim and I really hope that these MGMC methods will be really helpful to calculate these exponents at the critical temperature.

References

[Amo74a] D. E. Amos, *Computation of modified Bessel functions and their ratios*, Math. Comp. **28** (1974), 239–251. MR0333287

[Amo74b] D. E Amos, *Computation of modified bessel functions and their ratios*, Mathematics of Computation **28** (1974), no. 125, 239–251.

[And58] P. W. Anderson, *Random-phase approximation in the theory of superconductivity*, Phys. Rev. (2) **112** (1958), 1900–1916. MR0101138

[AVZ00] N. Angelescu, A. Verbeure, and V. A. Zagrebnov, *Quantum n-vector anharmonic crystal. II. Displacement fluctuations*, J. Statist. Phys. **100** (2000), no. 5-6, 829–851. MR1798545

[AZ85] N. Angelescu and V. A. Zagrebnov, *A generalized quasiaverage approach to the description of the limit states of the n-vector Curie-Weiss ferromagnet*, J. Statist. Phys. **41** (1985), no. 1-2, 323–334. MR813024

[BGR94] A Brandt, M Galun, and D Ron, *Optimal multigrid algorithms for calculating thermodynamic limits*, Journal of Statistical Physics **74** (1994), no. 1-2, 313–348.

[Bru67] S. G Brush, *History of the lenz-ising model*, Reviews of modern physics **39** (1967), no. 4, 883–893.

[BTW87] P. Bak, C. Tang, and K. Wiesenfeld, *Self-organized criticality: An explanation of the 1/f noise*, Physical review letters **59** (1987), no. 4, 381–384.

[CS11] S. Chatterjee and Q.-M. Shao, *Nonnormal approximation by Stein's method of exchangeable pairs with application to the Curie-Weiss model*, Ann. Appl. Probab. **21** (2011), no. 2, 464–483. MR2807964

[CT06] T. M. Cover and J. A. Thomas, *Elements of information theory*, Second, Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, 2006. MR2239987

[EHT00] R. S. Ellis, K. Haven, and B. Turkington, *Large deviation principles and complete equivalence and nonequivalence results for pure and mixed ensembles*, J. Statist. Phys. **101** (2000), no. 5-6, 999–1064. MR1806714

[Ell06] R. S. Ellis, *Entropy, large deviations, and statistical mechanics*, Classics in Mathematics, Springer-Verlag, Berlin, 2006. Reprint of the 1985 original. MR2189669

[EN78a] R. S. Ellis and C. M. Newman, *Limit theorems for sums of dependent random variables occurring in statistical mechanics*, Z. Wahrsch. Verw. Gebiete **44** (1978), no. 2, 117–139. MR0503333

[EN78b] R. S. Ellis and C. M. Newman, *The statistics of Curie-Weiss models*, J. Statist. Phys. **19** (1978), no. 2, 149–161. MR0503332

[Gri64] R. B. Griffiths, *Peierls proof of spontaneous magnetization in a two-dimensional ising ferromagnet*, Phys. Rev. **136** (1964Oct), A437–A439.

[IS90] E. K. Ifantis and P. D. Siafarikas, *Inequalities involving Bessel and modified Bessel functions*, J. Math. Anal. Appl. **147** (1990), no. 1, 214–227. MR1044696

[Isi25] E. Ising, *Contribution to the Theory of Ferromagnetism*, Z. Phys. **31** (1925), 253–258.

[KDB89] D. Kandel, E. Domany, and A. Brandt, *Simulations without critical slowing down: Ising and three-state potts models*, Physical Review B **40** (1989), no. 1, 330.

[KDR+88] D. Kandel, E. Domany, D. Ron, A. Brandt, and E. Loh Jr, *Simulations without critical slowing down*, Physical review letters **60** (1988), no. 16, 1591.

[KM13] K. Kirkpatrick and E. Meckes, *Asymptotics of the mean-field Heisenberg model*, J. Stat. Phys. **152** (2013), no. 1, 54–92. MR3067076

[KN16a] K. Kirkpatrick and T Nawaz, *Asymptotics of mean-field o(n) models*, J Stat Phys **65** (2016), 1114–1140.

[KN16b] K. Kirkpatrick and T. Nawaz, *Critical behavior of mean-field xy and related models*, Birkhauser Proceedings of the Conference in honor of Rodrigo Bauelos. (2016).

[KS89] H. Kesten and R. H. Schonmann, *Behavior in large dimensions of the Potts and Heisenberg models*, Rev. Math. Phys. **1** (1989), no. 2-3, 147–182. MR1070088

[KT73] J. M. Kosterlitz and D. J. Thouless, *Ordering, metastability and phase transitions in two-dimensional systems*, Journal of Physics C: Solid State Physics **6** (1973), no. 7, 1181.

[Mec09] E. Meckes, *On Stein's method for multivariate normal approximation*, High dimensional probability V: the Luminy volume, 2009, pp. 153–178. MR2797946

[Moo69] M. Moore, *Additional evidence for a phase transition in the plane-rotator and classical heisenberg models for two-dimensional lattices*, Physical Review Letters **23** (1969), no. 15, 861.

[MW66] N. D. Mermin and H. Wagner, *Absence of ferromagnetism or antiferromagnetism in one- or two-dimensional isotropic heisenberg models*, Phys. Rev. Lett. **17** (1966Nov), 1133–1136.

[Ron90] D. Ron, *Developments of fast numerical solvers for problems in optimization and statistical mechanics*, 1990.

[Ros17] L. Ross, *Dynamics of the mean-field heisenberg model.*, PhD dissertation, 2017.

[SDHR04] C. Stein, P. Diaconis, S. Holmes, and G. Reinert, *Use of exchangeable pairs in the analysis of simulations*, Stein's method: expository lectures and applications, 2004, pp. 1–26. MR2118600

[SS84] H. C. Simpson and S. J. Spector, *Some monotonicity results for ratios of modified Bessel functions*, Quart. Appl. Math. **42** (1984), no. 1, 95–98. MR736509

[Sta68] H E. Stanley, *Dependence of critical properties on dimensionality of spins*, Physical Review Letters **20** (1968), no. 12, 589–592.

[Ste86] C. Stein, *Approximate computation of expectations*, Institute of Mathematical Statistics Lecture Notes—Monograph Series, 7, Institute of Mathematical Statistics, Hayward, CA, 1986. MR882007

[VZ94] A. Verbeure and V. Zagrebnov, *Gaussian, non-gaussian critical fluctuations in the curie-weiss model*, Journal of statistical physics **75** (1994), no. 5-6, 1137–1152.